Benchmark Selection Guide, Vol. 1

*Doing it Right*

Art Morgan

Performance Engineer, Intel Corporation

Spring 2015
Legal Disclaimers

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark* and MobileMark*, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to http://www.intel.com/performance.

Intel is a sponsor and member of the BenchmarkXPRT* Development Community, and was the major developer of the XPRT* family of benchmarks. Principled Technologies* is the publisher of the XPRT family of benchmarks. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases.

© 2015 Intel Corporation. Intel, Intel Inside and other logos are trademarks or registered trademarks of Intel Corporation, or its subsidiaries in the United States and other countries.

* Other names and brands may be claimed as the property of others.
About this Guide

This guide uses benchmark comparison tables to help you select the right set of tools for Windows*, Android* and Chrome* platform evaluations.

- General Guidance on Platform Evaluations
- Benchmark Selection for Windows
- Benchmark Selection for Android
- Benchmark Selection for Chrome
- Basic Benchmark Selection Criteria
- Recommended Benchmarks by Form Factor
- Wrap Up

**Benchmarks covered in this guide:**
- 3DMark* 1.2.0
- ANDEBench* PRO
- AnTuTu* v5.6.1
- BaseMark* CL
- BatteryXPRT* 2014
- CompuBench* CL 1.5 for OpenCL*
- CompuBench* RS 2.0 for RenderScript*
- CrXPRT* 2015
- Geekbench* 3.3.1
- GFXBench* 3.0
- MobileXPRT* 2013
- Octane* 2.0
- PassMark* Performance Test 8.0
- PCMark* 8 v2.0 – Accelerated
- PCMark* 8 v2.0 – Conventional
- RoboHornet* RH-A1
- SunSpider* 1.0.2
- SYSmark* 2014
- TabletMark* v3
- TouchXPRT* 2014
- WebGL* Aquarium
- WebXPRT* 2013
How to Use this Guide

**Step ①**: Learn about Benchmark Domains and compare benchmarks within each domain

**Step ②**: Use comparison tables to select objective benchmarks in each Benchmark Domain

**Step ③**: Perform comprehensive and balanced platform evaluations using selected set of benchmarks
"Our position is that the only consistent and reliable measure of performance is the execution time of real programs, and that all proposed alternatives to time as the metric or to real programs as the items measured have eventually led to misleading claims or even mistakes in computer design."

General Guidance on Platform Evaluations

Translate technical arguments into business statements by specifying requirements in terms of end user benefits.

- Avoid using clock frequency or core count to specify platform performance
- Avoid using artificial or synthetic benchmarks – they do not represent real application behavior
- Avoid a sole “all purpose” benchmark for platform evaluations
Reporting Performance

“The guiding principle of reporting performance measurements should be **reproducibility** – list everything another experimenter would need to duplicate the results.”

There’s no such thing as a single “all purpose” benchmarking tool. You have to pick the right set of tools for the right jobs. Different sets of benchmarks are used for each type of benchmarking activity.

- Design & Development – specific tests
- Casual Tinkering – tire kicking
- Platform Evaluations – reproducible methodology; objective tests
“To overcome the danger of placing too many eggs in one basket, collections of benchmark applications, called benchmark suites, are a popular measure of performance of processors with a variety of applications. Of course, such suites are only as good as the constituent individual benchmarks. Nonetheless, a key advantage of such suites is that the weakness of any one benchmark is lessened by the presence of the other benchmarks.”

Different benchmarks test different areas of platform performance. For comprehensive platform evaluations, it’s best to cover all three domains of performance.

- Mainstream Apps – use standard OS services
- Web Apps – use web browser technologies; inherently cross-OS
- Game & GPGPU Apps – use graphics
Mainstream Focus

“Perhaps the most important and pervasive principle of computer design is to focus on the common case: In making a design trade-off, favor the frequent case over the infrequent case. This principle applies when determining how to spend resources, since the impact of the improvement is higher if the occurrence is frequent.”

## How to Read a Benchmark Comparison Table

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>What the Characteristic Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage Model</td>
<td>Short description of the usage scenario being tested by the benchmark. Not applicable if the benchmark is artificial or synthetic.</td>
</tr>
<tr>
<td>Publisher</td>
<td>Organization that designs, develops and/or distributes the benchmark.</td>
</tr>
<tr>
<td>Publisher Organization</td>
<td>Types of organizations include Benchmark Consortium, Benchmark Community (or Open Source Community), Independent Benchmark Vendor, Independent Software Vendor or Small Company.</td>
</tr>
<tr>
<td>Publisher Founded &gt; 10 Years</td>
<td>Benchmark publisher was founded at least 10 years ago.</td>
</tr>
<tr>
<td>Real-World Usages</td>
<td>Benchmark uses real applications or benchmark applications executing real workloads, and be based on real-world scenarios and workflows. All objective benchmarks are application-based or based on application code.</td>
</tr>
<tr>
<td>Open Development Process</td>
<td>The benchmark was designed with stakeholder input baked into the development process, guided by industry best practices and transparency.</td>
</tr>
<tr>
<td>Documented Scoring Methodology</td>
<td>Documentation is available that describes how a score is calculated, along with contributed weights for each subtest (if applicable) score.</td>
</tr>
<tr>
<td>Stable Release History</td>
<td>Benchmark is stable and test results are consistent from one release to the next. Any variation falls within nominal engineering tolerances.</td>
</tr>
</tbody>
</table>
Benchmark Selection for Windows*

Including a Discussion on GPGPU Computing
Windows* Benchmark Domains

Create a set of benchmarking tools by selecting a benchmark in each domain to get to a comprehensive view of platform performance.

- Mainstream Apps – execute standard libraries, system APIs and OS calls
- Web Apps – primarily JavaScript* and HTML5; sensitive to browser
- Game & GPGPU Apps – execute DirectX*, OpenGL* or OpenCL*
## Mainstream App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SYSmark* 2014</th>
<th>TouchXPRT* 2014</th>
<th>TabletMark* v3</th>
<th>PCMark* 8 v2.0 – Conventional</th>
<th>PassMark* Performance Test 8.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>Windows* Desktop Productivity</td>
<td>Windows Light Media Editing</td>
<td>Cross-Platform Light Productivity &amp; BL</td>
<td>Windows Everyday Computing</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>BAPCo*</td>
<td>Principled Technologies*</td>
<td>BAPCo</td>
<td>Futuremark*</td>
<td>PassMark Software</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Benchmark Consortium</td>
<td>Benchmark Community</td>
<td>Benchmark Consortium</td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
</tr>
<tr>
<td><strong>Publisher Founded &gt; 10 Years</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Real-World Usages</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td><strong>Open Development Process</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Documented Scoring Methodology</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>** Stable Release History**</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

- ✓: Partial characteristic – such as limited source code access or mostly synthetic.
# Web App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>WebXPRT* 2013</th>
<th>Octane* 2.0</th>
<th>SunSpider* 1.0.2</th>
<th>WebGL* Aquarium</th>
<th>RoboHornet* RH-A1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>Web Applications on Specified Browser</td>
<td>Synthetic</td>
<td>Synthetic</td>
<td>Synthetic</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Principled Technologies*</td>
<td>Google*</td>
<td>WebKit*</td>
<td>Human Engines*</td>
<td>RoboHornet GitHub*</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Benchmark Community</td>
<td>Open Source Community</td>
<td>Open Source Community</td>
<td>Small Company</td>
<td>Open Source Community</td>
</tr>
<tr>
<td>Publisher Founded &gt; 10 Years</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Real-World Usages</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Open Development Process</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Documented Scoring Methodology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Stable Release History</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ Partial characteristic – such as limited source code access or mostly synthetic.

---

**Basic Criteria Recommendations**

- **Web Application on Specified Browser**: Synthetic
- **Publisher**: Principled Technologies*
- **Publisher Organization**: Benchmark Community
- **Publisher Founded > 10 Years**: ✓
- **Real-World Usages**: ✓
- **Open Development Process**: ✓
- **Documented Scoring Methodology**: ✓
- **Stable Release History**: ✓
# Game & GPGPU App Benchmarks

## Basic Criteria Recommendations

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>3DMark* 1.2.0</th>
<th>GFXBench* 3.0</th>
<th>CompuBench* CL 1.5 Desktop</th>
<th>PCMark* 8 v2.0 – Accelerated</th>
<th>BaseMark* CL 1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>DirectX* / OpenGL* ES Gaming</td>
<td>DirectX / OpenGL ES Gaming</td>
<td>Synthetic</td>
<td>OpenCL* GPGPU Applications</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Futuremark*</td>
<td>Kishonti Informatics*</td>
<td>Kishonti Informatics</td>
<td>Futuremark</td>
<td>Rightware*</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
</tr>
<tr>
<td><strong>Publisher Founded &gt; 10 Years</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Real-World Usages</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Open Development Process</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Documented Scoring Methodology</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Stable Release History</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ Partial characteristic – such as limited source code access or mostly synthetic.
Spotlight: SYSmark*

Profile

Current Version & OS: SYSmark 2014 for Desktop Windows*
Usage Model: Windows Desktop Productivity
Publisher: BAPCo*
Information: Whitepaper, Results Database
Download From: BAPCo Online Store
Sideload From: N/A
Google Play* Installs: N/A
Pricing: $1,295 (SYSmark 2014 Medium Business)
Approximate Footprint: 25 GB
Approximate Running Time: 1 hour
Technologies Tested: HDR Imaging and Photo Stitching

Usage Scenarios
- Office Productivity
- Media Creation
- Data / Financial Analysis

What’s New & Interesting
- New compact usage scenarios
- New and updated business applications and workloads:
  - Microsoft OneNote*
  - Google Chrome*
  - Trimble SketchUp*
- Support for both Windows 32-bit and 64-bit in a single installation
- Easier to install, with a running time of less than 1 hour
## SYSmrk* 2014 Applications & Workloads

### Office Productivity
- Microsoft Word* 2013
- Microsoft Excel* 2013
- Microsoft PowerPoint* 2013
- Microsoft Outlook* 2013
- Microsoft OneNote* 2013
- Google Chrome*
- Adobe Acrobat* XI Pro
- WinZip* 17.5

### Media Creation
- Adobe Photoshop* CS6
- Adobe Premiere* Pro CS6 (for Windows* 64-bit)
- Adobe Premiere Pro CS4 (for Windows 32-bit)
- Trimble SketchUp* Pro 2013

### Data / Financial Analysis
- Microsoft Excel 2013
- WinZip 17.5

### Notes on Enhanced Workloads
- **Word**: Same mail merge workload as 2012, new PDF conversion
- **Outlook**: New workflow using 2012 dataset
- **Acrobat**: Two workloads same as 2012, one new document set
- **Chrome**: Browses web pages used by TabletMark* benchmark
- **Photoshop**: New HDR imaging and Photo Stitching workloads
- **Sketchup**: New workload based on Google* Sketchup version in 2012

### New Application
- New Workload
- Enhanced Workload
Information on GPGPU Computing

What you need to know about GPGPU Computing ...

- Good for predictable, uniform tasks; requires parallelization – **not good for mainstream** Windows* applications
- Constrained by Amdahl's Law: speedup of a program limited by the time needed for the sequential fraction
- Today, mainstream applications like Microsoft Office* benefit mostly from fewer, more powerful cores

Discussion

Also known as Heterogeneous System Architecture* (HSA), OpenCL*, CUDA*, Stream*, DirectCompute* and RenderScript* (on Android*). Benefits predictable, uniform tasks and requires the parallelization of application code to be effective, and generally does not benefit mainstream Windows applications like Microsoft Office. That's because GPGPU Computing is limited by Amdahl's Law: The speedup of a program using multiple processors in parallel computing is limited by the time needed for the sequential fraction of the program. **Today, mainstream applications have large sections of code that must remain sequential** – fewer, more powerful cores are preferred.

Profile

Current Version & OS: PCMark 8 v2.0 for Desktop Windows*
Usage Model: Windows Everyday Computing
Publisher: Futuremark*
Information: FAQ, Whitepaper, Results Database
Download From: Futuremark Website
Sideload From: N/A
Google Play* Installs: N/A
Pricing: $1,495 (PCMark 8 Professional Edition)
Approximate Footprint: 7 GB
Approximate Running Time: 30 minutes
Technologies Tested: OpenCL*

Hardware Tests
- PCMark 8 Home – Conventional / Accelerated
- PCMark 8 Creative – Conventional / Accelerated
- PCMark 8 Work – Conventional / Accelerated
- PCMark 8 Storage
- PCMark 8 Applications – Adobe CS*6
- PCMark 8 Applications – Microsoft Office*

Note: Scores are not compatible with Android* version.

Benchmark Assessment Disclosure

We currently do not recommend v2.0.191 released on January 17, 2014. Our current assessment is:

- PCMark 8 does not generate an overall "PCMark" score. This can lead to potentially misleading marketing of a specific subtest versus looking at the overall performance of a processor using the entire suite of PCMark 8 subtests.

- With the latest release, Futuremark enabled heavy OpenCL workloads on all of the Basic Edition subtests in what it calls "accelerated" mode, transforming PCMark 8 into an OpenCL (GPGPU) benchmark.

- Futuremark has released several new versions of PCMark 8 since its launch. Even Futuremark warns the user to be careful comparing results from different versions.
Maslow’s Hammer – Everything Looks Like a Nail

Common thread is showcasing GPU performance
GPGPU Effects are not Mainstream Usages
PassMark* Performance Test

Profile
Current Version & OS: PassMark V8.0 for Desktop Windows*
Usage Model: Synthetic
Publisher: PassMark Software
Information: Results Database, Version History
Download From: PassMark Website
Sideload From: N/A
Google Play* Installs: N/A
Pricing: $27
Approximate Footprint: 100 MB
Approximate Running Time: N/A
Technologies Tested: 2D Graphics, DirectX* 3D Graphics

Hardware Tests
- CPU Tests
- 2D Graphics Tests
- 3D Graphics Tests
- Disk Tests
- Memory Tests

Note: Scores are not compatible with Android* & iOS* versions.

Benchmark Assessment Disclosure
We currently do not recommend PassMark Performance Test. Our current assessment is:
- Since 2001, PassMark Software hosts a consumer-oriented “CPU Mark” results database at cpubenchmark.net. The results are posted by PassMark users with no system configuration information and can change from day-to-day because they are averaged.
- “CPU Mark” unrealistically scales with the number of cores. All component tests except the last use all available processor cores. No mainstream desktop application behaves this way.

PassMark “CPU Mark”
- Integer Math (high weight in score)
- Floating Point Math
- Prime Numbers
- Extended Instructions (SSE)
- Compression (high weight in score)
- Encryption
- Physics
- Sorting
- Single Threaded

* indicates mobile app platform compatibility.
Data from a Consumer Website is Unreliable

- Website: Free, Amateur, GarbageInGarbageOut, Transient, Abundant, Consumer, Changeable, Amorphous, Averaged, Uncertain, Unreliable, Indefinite, Variable
- Lab: Business, Defined, Stable, Commercial, Limited, Reliable, Repeatable, Rigorous, Trusted, Configurations, Deterministic
Benchmark Selection for Android*

Including a Discussion on Multi-Core Computing
Android* Benchmark Domains

Create a set of benchmarking tools by selecting a benchmark in each domain to get to a comprehensive view of platform performance.

- Mainstream Apps – execute standard libraries, system APIs and OS calls
- Web Apps – primarily JavaScript* and HTML5; sensitive to browser
- Game & GPGPU Apps – execute OpenGL* or RenderScript*
# Mainstream App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>MobileXPRT* 2013</th>
<th>TabletMark* v3</th>
<th>ANDEBench* PRO</th>
<th>AnTuTu* v5.6.1</th>
<th>Geekbench* 3.3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>Android* Light Media Editing</td>
<td>Cross-Platform Light Productivity &amp; BL</td>
<td>Android Processor, Workflow &amp; 3D Gfx</td>
<td>Synthetic</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Principled Technologies*</td>
<td>BAPCo*</td>
<td>EEMBC*</td>
<td>AnTuTu Labs</td>
<td>Primate Labs*</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Benchmark Community</td>
<td>Benchmark Consortium</td>
<td>Benchmark Consortium</td>
<td>Small Company</td>
<td>Small Company</td>
</tr>
<tr>
<td><strong>Publisher Founded &gt; 10 Years</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Real-World Usages</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Open Development Process</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Documented Scoring Methodology</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Stable Release History</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ Partial characteristic – such as limited source code access or mostly synthetic.
# Web App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>WebXPRT* 2013</th>
<th>Octane* 2.0</th>
<th>SunSpider* 1.0.2</th>
<th>WebGL* Aquarium</th>
<th>RoboHornet* RH-A1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>Web Applications on Specified Browser</td>
<td>Synthetic</td>
<td>Synthetic</td>
<td>Synthetic</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Principled Technologies*</td>
<td>Google*</td>
<td>WebKit*</td>
<td>Human Engines*</td>
<td>RoboHornet GitHub*</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Benchmark Community</td>
<td>Open Source Community</td>
<td>Open Source Community</td>
<td>Small Company</td>
<td>Open Source Community</td>
</tr>
<tr>
<td><strong>Publisher Founded &gt; 10 Years</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Real-World Usages</strong></td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Open Development Process</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Documented Scoring Methodology</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Stable Release History</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ Partial characteristic – such as limited source code access or mostly synthetic.

---

**Basic Criteria Recommendations**

- **Web Applications on Specified Browser**
- **Synthetic**
- **WebKit**
- **Open Source Community**
- **Small Company**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**
- **Open Source Community**

---

Intel® Processors

[Image: Intel logo]
### Game & GPGPU App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>3DMark* 1.2.0</th>
<th>GFXBench* 3.0</th>
<th>CompuBench* RS 2.0 Mobile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usage Model</strong></td>
<td>OpenGL* ES 2.0 Gaming</td>
<td>OpenGL ES 3.0 Gaming</td>
<td>Synthetic</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Futuremark*</td>
<td>Kishonti Informatics*</td>
<td>Kishonti Informatics</td>
</tr>
<tr>
<td><strong>Publisher Organization</strong></td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
<td>Independent Benchmark Vendor</td>
</tr>
<tr>
<td><strong>Publisher Founded &gt; 10 Years</strong></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
</tr>
<tr>
<td><strong>Real-World Usages</strong></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
</tr>
<tr>
<td><strong>Open Development Process</strong></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
</tr>
<tr>
<td><strong>Documented Scoring Methodology</strong></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
</tr>
<tr>
<td><strong>Stable Release History</strong></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
<td><img src="checkmark.png" alt="Checkmark" /></td>
</tr>
</tbody>
</table>

- Partial characteristic – such as limited source code access or mostly synthetic.
Spotlight: MobileXPRT*

Profile
Current Version & OS: MobileXPRT 2013 for Android*
Usage Model: Android Light Media Editing
Publisher: Principled Technologies*
Information: FAQ, Community Open Source
Download From: Google Play*
Sideload From: Principled Technologies Website
Google Play Installs: 5,000 – 10,000
Pricing: Free
Approximate Footprint: 650 MB
Approximate Running Time: 10 minutes (Performance Tests)
Technologies Tested: Runtime and Native Development Kit (NDK)

Usage Scenarios – Performance Tests
- Apply Photo Effects
- Create Photo Collages
- Create Slideshow
- Encrypt Personal Content
- Detect Faces in Photos

Usage Scenarios – UX Tests
- List Scroll
- Grid Scroll
- Gallery Scroll
- Browser Scroll
- Zoom & Pinch
Information on Multi-Core Computing

What you need to know about Multi-Core Computing ...

- Good for predictable, uniform tasks; requires parallelization – not good for mainstream Android* applications
- Constrained by Amdahl’s Law: the performance becomes dependent on the performance in the sequential parts
- Today, Android apps like Facebook*, Twitter* and Chrome* benefit mostly from fewer, more powerful cores

Discussion

Multitasking or OS-level process scheduling aside, Multi-Core Computing requires the parallelization of application code, and generally does not benefit mainstream Android applications like Facebook, Twitter and Chrome. That’s because Multi-Core Computing is limited by a corollary to Amdahl’s Law: The more cores you add to a CPU, the faster the parallel parts of an application are processed, so the more the performance becomes dependent on the performance in the sequential parts. Today, the sequential code sections of mainstream applications are the long pole in the tent, and are better serviced by fewer, more powerful cores.

AnTuTu*

Profile
Current Version & OS: AnTuTu v5.6.1 for Android*
Usage Model: Synthetic
Publisher: AnTuTu Labs
Information: Results Database
Download From: Google Play*
Sideload From: N/A
Google Play Installs: 10,000,000 – 50,000,000
Pricing: Free
Approximate Footprint: 55 MB
Approximate Running Time: 4 minutes
Technologies Tested: Runtime and OpenGL* ES

Benchmark Assessment Disclosure
We currently do not recommend AnTuTu. Our assessment is:

- The CPU test is believed to be based on NBench*, authored by BYTE* magazine back in the 1980’s. AnTuTu artificially utilizes all available execution threads and is not representative of a real Android application. A higher AnTuTu score does not necessary mean a better user experience.

- There is little publicly available information about exactly what AnTuTu does or how it does it. Because of the simplistic nature of the benchmark, it is susceptible to what the industry is calling "benchmark boosting".

- With new releases on an unpredictable schedule, runs from one week to the next cannot be reliably compared.

- Ask yourself this question: Other than higher is better, do you know what an AnTuTu score actually means?

Hardware Tests
- UX – Multitask, Runtime
- CPU (multithreaded) – Integer, Float-Point [sic]
- CPU (single threaded) – Integer, Float-Point [sic]
- RAM – RAM Operation, RAM Speed
- GPU – 2D Graphics, 3D Graphics
- I/O – Storage I/O, Database I/O
The Objective Benchmark Difference

Benchmark metrics in the example below help to illustrate that MobileXPRT* is based on real usage scenarios, while AnTuTu* is based on component tests.

<table>
<thead>
<tr>
<th>MobileXPRT 2013 Metrics</th>
<th>AnTuTu v5.6.1 Metrics</th>
<th>Hardware Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apply Photo Effects</td>
<td>UX Multitask</td>
<td></td>
</tr>
<tr>
<td>Create Photo Collages</td>
<td>UX Runtime</td>
<td></td>
</tr>
<tr>
<td>Create Slideshow</td>
<td>CPU (st/mt) Integer</td>
<td></td>
</tr>
<tr>
<td>Encrypt Personal Content</td>
<td>CPU (st/mt) Float [sic]</td>
<td></td>
</tr>
<tr>
<td>Detect Faces in Photos</td>
<td>RAM Operation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAM Speed</td>
<td></td>
</tr>
<tr>
<td><strong>User Experience</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>List Scroll</td>
<td>GPU 2D Graphics</td>
<td></td>
</tr>
<tr>
<td>Grid Scroll</td>
<td>GPU 3D Graphics</td>
<td></td>
</tr>
<tr>
<td>Gallery Scroll</td>
<td>I/O Storage</td>
<td></td>
</tr>
<tr>
<td>Browser Scroll</td>
<td>I/O Database</td>
<td></td>
</tr>
<tr>
<td>Zoom and Pinch</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Benchmark Selection for Chrome*

Including a Discussion on Chrome Computing
Chrome* Benchmark Domains (Web Apps)

Chrome applications run in the browser, therefore benchmarks for platform evaluations are limited to Web Apps.

- Mainstream Apps – run in Chrome browser
- Web Apps – primarily JavaScript*, HTML5 and WebGL*
- Game & GPGPU Apps – run in Chrome browser using WebGL
## Web App Benchmarks

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>WebXPRT* 2013</th>
<th>CrXPRT* 2015</th>
<th>Octane* 2.0</th>
<th>WebGL* Aquarium</th>
<th>RoboHornet* RH-A1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage Model</td>
<td>Web Applications on Specified Browser</td>
<td>Chromebook* Perf Qualified Battery Life</td>
<td>Synthetic</td>
<td>Synthetic</td>
<td>Synthetic</td>
</tr>
<tr>
<td>Publisher</td>
<td>Principled Technologies*</td>
<td>Principled Technologies</td>
<td>Google*</td>
<td>Human Engines*</td>
<td>RoboHornet GitHub*</td>
</tr>
<tr>
<td>Publisher Organization</td>
<td>Benchmark Community</td>
<td>Benchmark Community</td>
<td>Open Source Community</td>
<td>Small Company</td>
<td>Open Source Community</td>
</tr>
<tr>
<td>Publisher Founded &gt; 10 Years</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Real-World Usages</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Open Development Process</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Documented Scoring Methodology</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>❌</td>
<td>✓</td>
</tr>
<tr>
<td>Stable Release History</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ Partial characteristic – such as limited source code access or mostly synthetic.
Spotlight: WebXPRT*

Profile

**Current Version & OS:** WebXPRT 2013 for All Major OS  
**Usage Model:** Web Applications on Specified Browser  
**Publisher:** Principled Technologies*  
**Information:** FAQ, Whitepaper, Community Open Source  
**Download From:** Run on Principled Technologies Website  
**Sideloard From:** N/A  
**Google Play* Installs:** N/A  
**Pricing:** Free  
**Approximate Footprint:** N/A  
**Approximate Running Time:** 15 minutes  
**Technologies Tested:** HTML5, JavaScript* and DOM

Usage Scenarios

- Photo Effects
- Face Detect
- Stock Dashboard
- Offline Notes

What’s New & Interesting

- Next generation WebXPRT coming soon
- New photo collages media workload
- New create sales reports / charts productivity workload
- New Approximate Running Time: 20 minutes
- New Technologies Tested: HTML5 Web Workers for JavaScript background (multi-threaded) execution
- Next generation WebXPRT usage scenarios:
  - Photo Editing
  - Photo Organization
  - My Dashboard
  - Productivity
  - Education
Information on Chrome* Computing

What you need to know about Chrome Computing ...

- Revolves around the cloud computing model, where computing is processed by servers and delivered online
- Client performance is key because the user experience is ultimately delivered by the client device
- Benchmarking focus is on Chrome browser technologies such as JavaScript*, HTML5, DOM, NaCl* and WebGL*

Discussion

Chrome Computing revolves around the cloud computing model, where applications and data are processed on servers and delivered online. For a Chrome client such as a Chromebook*, performance is key because the user experience is ultimately delivered by the hardware and software of the client device. Chrome clients execute Chrome OS, a barebones version of Linux* optimized for rapid booting and running the Chrome web browser. Lightweight, synthetic benchmarks like Dhrystone* can be ported to Chrome OS, however the benchmarking focus is on web browser technologies such as JavaScript, HTML5, DOM, NaCl (Google Native Client) and WebGL.
RoboHornet*

Profile
Current Version & OS: RoboHornet RH-A1
Usage Model: Synthetic
Publisher: RoboHornet GitHub*
Information: FAQ, Open Source
Download From: Run on RoboHornet Website
Sideload From: N/A
Google Play* Installs: N/A
Pricing: Free
Approximate Footprint: N/A
Approximate Running Time: 5 minutes
Technologies Tested: HTML5, JavaScript*, DOM, CSS Selectors

Hardware Tests
- Table Tests
- 2D Canvas Tests
- Scrolling Tests
- DOM Test
- Local Storage Tests

Benchmark Assessment Disclosure
We currently do not recommend RoboHornet. Our current assessment is:

- The major web browser technologies available today include HTML5, JavaScript, DOM and WebGL*. The tests in WebXPRT*, Octane*, SunSpider* and WebGL Aquarium* are representative of the major web browser technologies.

- Running additional browser benchmarks such as RoboHornet, Kraken*, Dramaeo*, Peacekeeper* and Browsermark*, which cover the same browser technologies, will show the same general performance characteristics.
Basic Benchmark Selection Criteria

For Objective Benchmarks
Basic Benchmark
Selection Criteria

Intel believes that good benchmarks should meet both of these two basic criteria:

1. The benchmark uses real applications or benchmark applications executing real workloads, and be based on real-world scenarios and workflows

2. The benchmark was designed with industry stakeholder input baked into the development process, guided by industry best practices and transparency
A modern benchmark should model actual user workflows such as creating a document, retouching a photo or editing a video. Also, a current trend is to include tests for advanced features such as UI gestures and security.

Intel's recommendation is to evaluate performance using benchmarks that represent real applications or include usage scenario applications from Independent Software Vendors.

If a simplistic hardware test is required, it is still preferred that the benchmarking tests actually execute real code and perform a diverse set of operations on real data.
Designed with Stakeholder Input

Designed with stakeholder input means that stakeholder input is baked into the development process. This is evident in established and trusted organizations such as BAPCo*, Principled Technologies*, SPEC*, EEMBC*, MobileBench*, Rightware* and Kishonti Informatics*

Depending on the organization, stakeholders can assist with coding, provide feedback and/or participate in pre-release testing

Transparency = reproducible results, whitepapers, sensitivity studies, full disclosure reports
Recommended Benchmarks by Form Factor

For Desktop, AIO, Notebook and Tablet
### Recommended Desktop Benchmarks

For comparing value to premium traditional Desktop platforms

<table>
<thead>
<tr>
<th>Benchmark Name</th>
<th>Usage Model Description</th>
<th>OS Support</th>
<th>Benchmark Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDXPRT* 2014</td>
<td>Windows* Media Editing</td>
<td></td>
<td>Principled Technologies*</td>
</tr>
<tr>
<td>TouchXPRT* 2014</td>
<td>Windows Light Media Editing</td>
<td></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>WebXPRT* 2013</td>
<td>Web Applications on Specified Browser</td>
<td></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>SYSmark* 2014</td>
<td>Windows Desktop Productivity</td>
<td></td>
<td>BAPCo*</td>
</tr>
<tr>
<td>SPEC* CPU2006</td>
<td>Compute Intensive Application Performance</td>
<td></td>
<td>SPEC</td>
</tr>
<tr>
<td>3DMark* Fire Strike 1.1.0</td>
<td>DirectX* 11 Gaming</td>
<td></td>
<td>Futuremark*</td>
</tr>
<tr>
<td>GFXBench* 3.0 Manhattan Offscreen</td>
<td>DirectX 10 / OpenGL* ES 3.0 Gaming</td>
<td></td>
<td>Kishonti Informatics*</td>
</tr>
</tbody>
</table>
## Recommended Notebook & AIO/pAIO Benchmarks

For comparing value to premium Notebook, Chromebook & AIO/pAIO platforms

<table>
<thead>
<tr>
<th>Benchmark Name</th>
<th>Usage Model Description</th>
<th>OS Support</th>
<th>Benchmark Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>CrXPRT* 2015</td>
<td>Chromebook* Performance Qualified Battery Life</td>
<td></td>
<td>Principled Technologies*</td>
</tr>
<tr>
<td>HDXPRT* 2014</td>
<td>Windows* Media Editing</td>
<td></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>TouchXPRT* 2014</td>
<td>Windows Light Media Editing</td>
<td></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>WebXPRT* 2013</td>
<td>Web Applications on Specified Browser</td>
<td></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>MobileMark* 2014</td>
<td>Windows Performance Qualified Battery Life</td>
<td></td>
<td>BAPCo*</td>
</tr>
<tr>
<td>SYSmark* 2014</td>
<td>Windows Desktop Productivity</td>
<td></td>
<td>BAPCo</td>
</tr>
<tr>
<td>TabletMark* v3</td>
<td>Cross-Platform Light Productivity, Media &amp; BL</td>
<td></td>
<td>BAPCo</td>
</tr>
<tr>
<td>SPEC* CPU2006</td>
<td>Compute Intensive Application Performance</td>
<td></td>
<td>SPEC</td>
</tr>
<tr>
<td>3DMark* Sky Diver 1.0.0 or Ice Storm 1.2.0 Unlimited</td>
<td>DirectX* / OpenGL* ES Gaming</td>
<td></td>
<td>Futuremark*</td>
</tr>
<tr>
<td>GFXBench* 3.0 Manhattan or T-Rex Offscreen</td>
<td>DirectX / OpenGL ES Gaming</td>
<td></td>
<td>Kishonti Informatics*</td>
</tr>
</tbody>
</table>
# Recommended Tablet & 2 in 1 Benchmarks

For comparing value to premium Tablet & 2 in 1 platforms

<table>
<thead>
<tr>
<th>Benchmark Name</th>
<th>Usage Model Description</th>
<th>OS Support</th>
<th>Benchmark Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANDEBench* PRO</td>
<td>Android* Processor, Workflow &amp; 3D Graphics</td>
<td><img src="android_icon.png" alt="Android" /></td>
<td>EEMBC*</td>
</tr>
<tr>
<td>BatteryXPRT* 2014</td>
<td>Android Performance Qualified Battery Life</td>
<td><img src="android_icon.png" alt="Android" /></td>
<td>Principled Technologies*</td>
</tr>
<tr>
<td>MobileXPRT* 2013</td>
<td>Android Light Media Editing</td>
<td><img src="android_icon.png" alt="Android" /></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>TouchXPRT* 2014</td>
<td>Windows* Light Media Editing</td>
<td><img src="windows_icon.png" alt="Windows" /></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>WebXPRT* 2013</td>
<td>Web Applications on Specified Browser</td>
<td><img src="android_icon.png" alt="Android" />, <img src="ios_icon.png" alt="iOS" /></td>
<td>Principled Technologies</td>
</tr>
<tr>
<td>MobileMark* 2014</td>
<td>Windows Performance Qualified Battery Life</td>
<td><img src="windows_icon.png" alt="Windows" /></td>
<td>BAPCo*</td>
</tr>
<tr>
<td>SYSmark* 2014</td>
<td>Windows Desktop Productivity</td>
<td><img src="windows_icon.png" alt="Windows" /></td>
<td>BAPCo</td>
</tr>
<tr>
<td>TabletMark* v3</td>
<td>Cross-Platform Light Productivity, Media &amp; BL</td>
<td><img src="android_icon.png" alt="Android" />, <img src="ios_icon.png" alt="iOS" /></td>
<td>BAPCo</td>
</tr>
<tr>
<td>SPEC* CPU2000†</td>
<td>Compute Intensive Application Performance</td>
<td><img src="windows_icon.png" alt="Windows" />, <img src="android_icon.png" alt="Android" />, <img src="ios_icon.png" alt="iOS" /></td>
<td>SPEC</td>
</tr>
<tr>
<td>3DMark* Ice Storm 1.2.0 Unlimited</td>
<td>DirectX* 9 / OpenGL* ES 2.0 Gaming</td>
<td><img src="android_icon.png" alt="Android" />, <img src="ios_icon.png" alt="iOS" /></td>
<td>Futuremark*</td>
</tr>
<tr>
<td>GFXBench* 3.0 T-Rex Offscreen</td>
<td>DirectX 9 / OpenGL ES 2.0 Gaming</td>
<td><img src="android_icon.png" alt="Android" />, <img src="ios_icon.png" alt="iOS" /></td>
<td>Kishonti Informatics*</td>
</tr>
</tbody>
</table>

† SPEC CPU2000 is supported on Windows* and UNIX*/Linux* platforms; however it is possible to have an unofficial port to other operating systems.
Wrap Up
Performance Strategy Workshop

Objectives: Align on platform performance goals, benchmarking methodologies, and map out a joint plan of action

Deliverables: Discussion notes and joint action plan

Attendees: Performance Engineering, Product Development, Product Management, Product Marketing, Product Procurement

Time Required: 2-3 hours

Workshop Agenda:

[10 min] Introductions

[15 min] Customer: Product performance strategy

[15 min] Intel: Platform benchmarking strategy

[15 min] Intel: Objective benchmark demos

[20 min] All: Vision of Success – Capture goals

[15 min] Intel: Platform evaluation framework

[20 min] All: Current Reality – Capture where we are today

[20 min] All: Joint Action Plan – Map out achieving goals

[10 min] All: Wrap Up & Next Steps

Contact Art.Morgan@intel.com for more information

* Other names and brands may be claimed as the property of others. Screenshots are simulated.
Summary &
Call to Action

Doing it Right: Create a set of benchmarking tools by selecting a benchmark in each domain to get to a comprehensive view of platform performance.

- Download and run all of the recommended benchmarks
- Schedule a Performance Strategy Workshop today
- Contact Art.Morgan@intel.com with any questions or feedback

* Other names and brands may be claimed as the property of others. Screenshots are simulated.