Do 8 Cores Really Matter in Smartphones?

A look at how CPU core count alone affect overall smartphone performance

Summary

In a recent live benchmarking session, Moor Insights & Strategy answered the question "Do 8 cores really matter in smartphones?" with a series of benchmarks and real application tests. This brief supports the video with performance and experience results.

As the smartphone industry has begun to mature, one-upmanship among smartphone manufacturers and SoC vendors has bred a dangerous trend: ever-increasing processor core counts and the association between increased CPU core count and greater performance. This association originated as SoC vendors and OEMs have tried to find ways to differentiate themselves from one another through core counts.

Some vendors are creating confusion, as phones today have core counts from 2 up to 8 and vary wildly in performance and, even more importantly, experience. One reason for this confusion is many users and reviewers have used inappropriate benchmarks to illustrate smartphone user experience and real world performance. As a result, we believe that some consumers are misled in their buying decisions and may end up with the wrong device and the wrong experience.

The 8 Core Myth

The 8 Core Myth, also known as the Octacore Myth, is the perception that more CPU cores are better and having more cores means higher performance. Today's smartphones range from 2 cores up to 8 cores, even though performance and user experience are not a function of CPU core count. The myth, however, will not be limited to 8 cores, as there are plans for SoCs with up to 10 cores, and we could even see more in the future.

Not All Cores Are the Same

In some phones, users are getting Octacore designs with up to 8 ARM Cortex-A53 cores. These 8 cores perform differently than 4 ARM Cortex-A57 cores paired with 4 ARM Cortex-A53 cores in what is called a big.LITTLE configuration. Core designs vary wildly from ARM's own A53 and A57 64-bit CPUs to Intel's x86 Atom 4-core processors to Apple's 2-core A8 ARM processor.

All these processors are designed differently and behave differently across application workloads and operating systems. Some cores are specifically designed for high performance, some for low power. Others are designed to balance the two through dynamic clocking and higher IPC (instructions per clock). As a result, no two SoCs necessarily perform the same when you take clock speed and core count into account.

This is inherently why **CPU core count alone should not be a factor in selecting a smartphone**.

ARM's 64-bit Cores Become Attractive

As Intel has pushed the industry toward 64-bits, many SoC vendors have adopted ARM's latest 64-bit processors including the A57 and A53. ARM's Cortex-A57 and A53 processors are the most common processors in mobile SoCs today from AllWinner, MediaTek, Nvidia, Qualcomm, and Samsung. Because all these vendors' use the same ARM cores, CPU performance across different SoC vendors has not varied much. So some vendors tout core count as a differentiator and stuff their SoCs with as many cores as they can. This ultimately leads to another problem: how to evaluate different SoCs against one another in benchmarks to determine the best performance.

The Mobile Benchmarking Problem

There are a lot of mobile benchmarks out there today; many of them are poorly written and also improperly used by reviewers. These two factors contribute to incorrect interpretations of benchmark results as a measure of what a user will experience. These misleading or misunderstood benchmarks and reviews ultimately hurt consumers and pull the industry in the wrong direction.

Making a Good Mobile Benchmark

The problem is that most benchmarks today do not reflect real application use—which is the entire point of benchmarking a smartphone.

A smartphone benchmark is supposed to quantify the **software experience** that a user can expect from the device based upon the tests contained within the benchmark and the results that they return. Benchmark results need to be **consistent**, providing the same result every time they are tested. Benchmarks need to run on **shipping devices**, and **one benchmark is never enough** to accurately depict a system's performance. Also, the companies that make the benchmarks need to be **transparent** about how they build their benchmarks and have a working relationship with SoC vendors. This includes monitoring and moderating any potential cheating by OEMs in order to boost their own devices' benchmark scores. This became an issue after technology blog Anandtech discovered and reported cheating across all of the top Android OEMs and forced the benchmark creators to address the issue and de-list some devices.

Good Benchmarking Practices

The best benchmarking practices take real world scenarios into account. This includes using shipping devices that consumers can buy; many reviewers get pre-production devices, but the better and more accurate reviews actually use shipping hardware. Another best practice is to consider what the benchmark is testing and whether that is enough to accurately assess the phone. Realistically, no single benchmark is enough to test a device, and it will take several benchmarks to evaluate a device's performance. Different benchmarks should test different parts of the system that might not be sufficiently covered by another benchmark, while still using real world style scenarios.

The Benchmarks, Apps, Tools, & Phones

The Benchmarks

MI&S ran three different benchmarks.

- **Futuremark PCMark for Android Work Performance v1.1**: While the "work" moniker is a bit confusing, it does separate it from the other Futuremark products. Futuremark is a respected benchmark company known for creating consistent, reliable benchmarks that properly serve their intended purpose. PCMark for Android is designed with the user in mind, combining **web browsing**, **video**, **writing**, and **photo editing** into one benchmark. All of these tests use real touch inputs and actual text and video to evaluate system performance in these "work" scenarios. However, this single benchmark is not enough to entirely evaluate a smartphone's performance, as it does not test 3D gaming or camera capabilities, which is why we are using three benchmarks.
- **Basemark X for Android v1.1**: A **3D gaming** benchmark specifically designed for mobile using Unity v4.2.2, one of the most commonly used engines in mobile gaming. It can switch between on-screen and off-screen testing, which serve different purposes. We chose on-screen testing as more representative of a user's experience rather than the theoretical performance at a certain resolution. To prevent devices from exceeding the refresh rate of the display, we ran the phones in the highest graphical setting; this ensured that we would not run into any frame rate limitations and would be experiencing the best possible graphics. We do respect that there are reasons to test off-screen, such as scenarios where only the part is tested or the device is connected to an external display.
- **Sofica CamSpeed for Android v2.31**: Designed to measure the overall **speed of a phone's camera**. It does not measure the quality of the image, as that can be very subjective, but rather sticks to an objective measurement of the camera system's overall speed. The overall speed of a camera in taking pictures is a huge factor for many smartphone users and as a result should be considered when benchmarking.

All three benchmarks were obtained directly through the benchmark creators and are press versions, which are more fully featured than their public versions. The benchmarks were selected for their availability as press versions as well as their availability on the Google Play Store. They were also selected for their consistency and openness towards how and what they tested without any features disabled.

The Apps

We ran two apps: **YouTube v10.24.57** and **WeChat v6.2**. We picked YouTube for its global popularity and amount of time mobile users spend on YouTube. We picked WeChat, because it is one of the most actively used apps in the world, especially China.

For both apps, 8 cores do not improve users' everyday experience. We disabled cores on some of the phones where possible, and showed how these applications are not slowed down by having fewer cores available.

The Tools
We used two tools.

- **Qualcomm Trepn Profiler**: Draws on a multitude of data points inside of an Android smartphone. We used it to measure clock speed and load of a smartphone while using WeChat and YouTube.
- **3C CPU Manager**: Part of the 3C Toolbox Pro application and requires the device to be rooted. Once rooted, it allows the user to adjust core clock speeds and to turn CPU cores on and off via hot plugging, if the device's kernel allows for it. Not all devices' kernels allowed for hot plugging, so only two devices had their cores disabled for testing.

The Phones
We tested five different smartphones. All had Android 5.0 except the HTC which had Android KitKat.

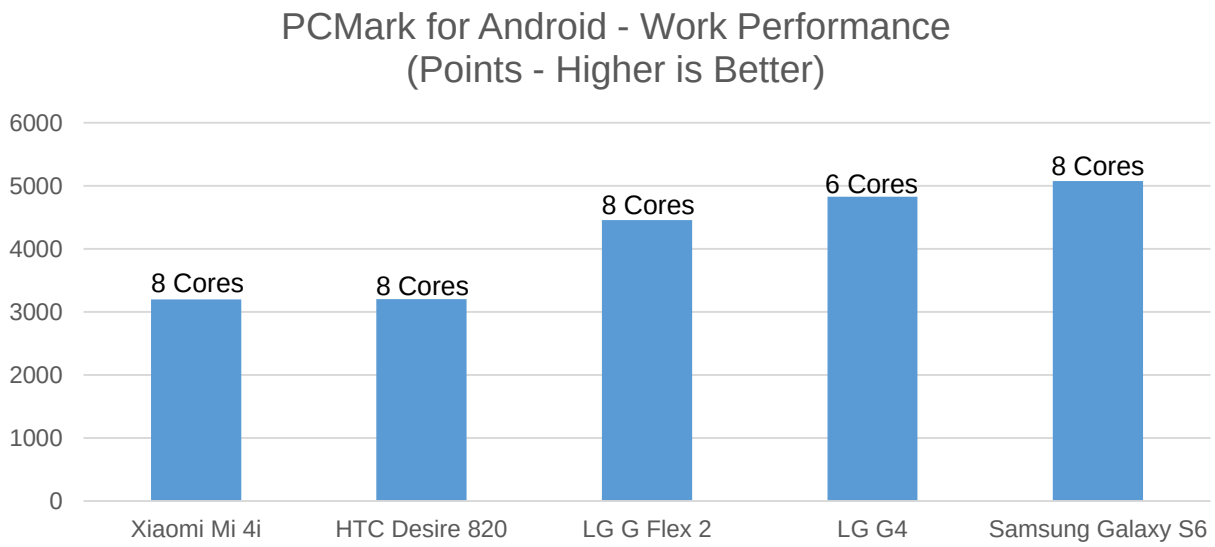
Table 1: Phones, SoCs, & Cores

Phone	SoC	Cores	Total Core Count
LG G4	Snapdragon 808	2x Cortex-A57 + 4x Cortex-A53	**6**
Samsung Galaxy S6	Exynos 7420	4x Cortex-A57 + 4x Cortex-A53	**8**
Xiaomi Mi 4i	Snapdragon 615	8x Cortex-A53	**8**
HTC Desire 820S	MediaTek MT6752	8x Cortex-A53	**8**
LG G Flex 2	Snapdragon 810	4x Cortex-A57 + 4x Cortex-A53	**8**

Benchmark Tests & Results
We tested the smartphones in the order above, using 3C CPU and Trepn along the way to disable cores as necessary and re-enabling them to prepare devices for the next test.

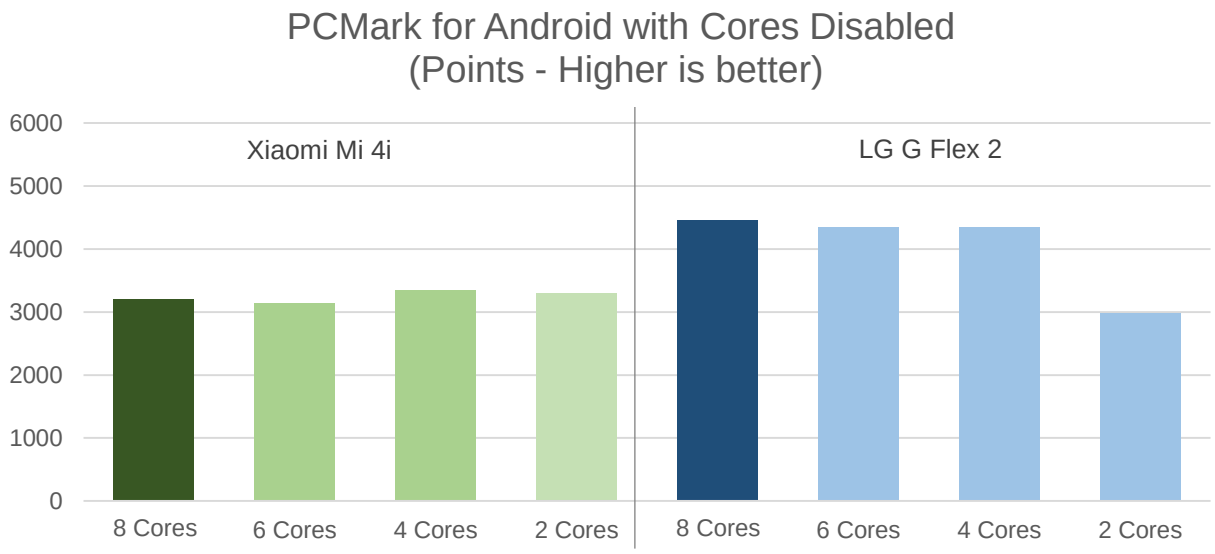
PCMark
In PCMark we did not initially disable any cores to show the wide breadth of scores from devices. Scores ranged from 3199 to 5078 with all of the phones except one (LG G4) being 8 core phones. Even though almost all of these phones are 8 core phones, they all perform differently from each other. Also, the 6-core LG G4 outperformed three of the 8-core phones and almost caught up to the Samsung Galaxy S6.

Figure 1: Same Number of Cores, Different Levels of Performance



We took the Xiaomi Mi 4i and LG G Flex 2 and disabled their cores from 8 down to 2 cores using the 3C CPU manager. PCMark performance was for the most part unaffected by CPU core count. The only huge loss was when the G Flex 2 was brought down to two A53 cores by disabling the last two A57s. Because the G Flex 2 has two different types of cores, we had to be sure we were disabling the right cores. Cores 0 through 3 were all A53s while Cores 4 through 7 were A57s. As a result, we first disabled Cores 7 and 6, followed by cores 2 and 3 and finally 4 and 5 to get down to 2 cores for the final test. Remember that both phones initially started with 8 cores.

Figure 2: Fewer Cores, Effect on Performance

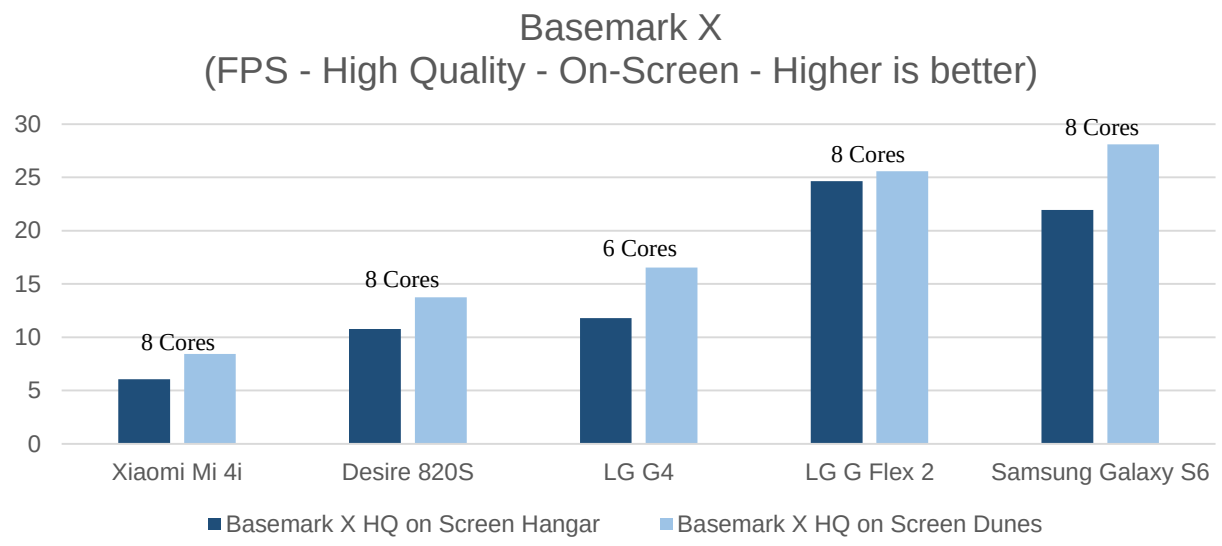


Xiaomi Mi 4i performance **went up** slightly when running only 4 or 2 cores versus 8 or 6. Performance for the G Flex 2 was relatively stable across 8, 6, and 4 CPU cores; only when it was down to the last two A53s did performance really decrease significantly.

Basemark X

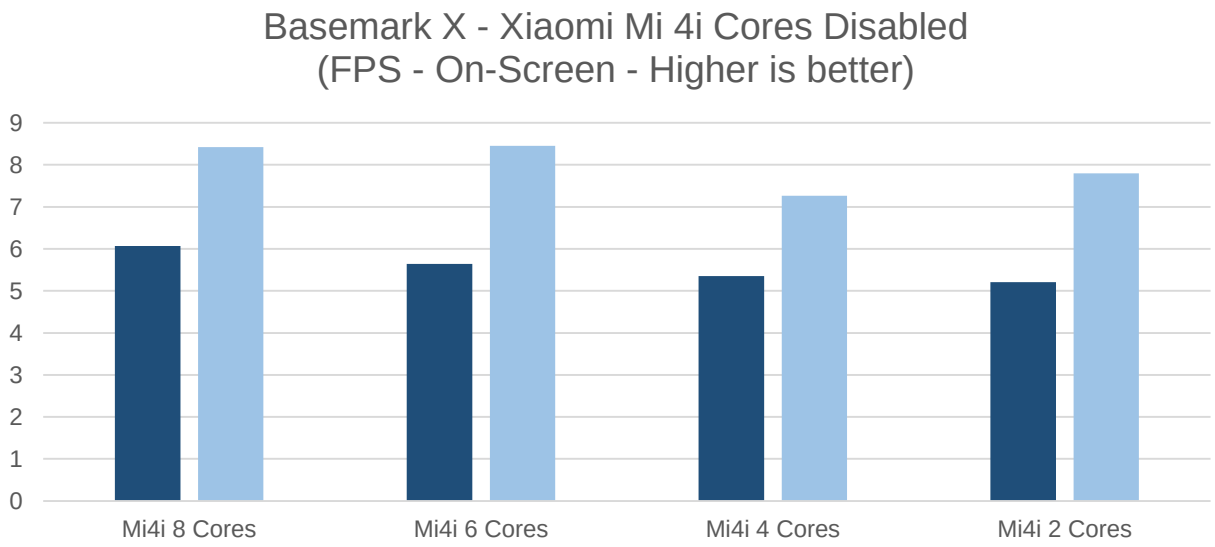
In Basemark X, the benchmark is made up of two different test scenes, so it generates two different average frame rate scores. We tested all devices with all cores enabled, and then we disabled cores in the same order before. Also keep in mind that we ran these tests at the highest settings and using on-screen resolution, so these scores will likely vary from what other reviewers have shown.

Figure 3: Frame Rates



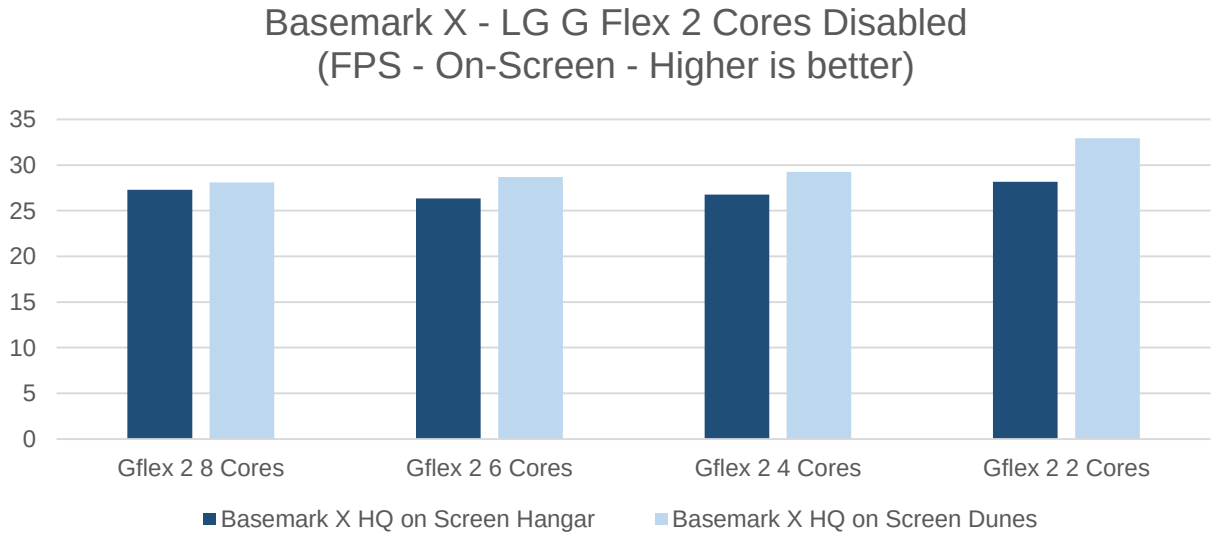
The Samsung Galaxy S6 and LG G Flex 2 had the highest frame rates in both scenes. However, the reality is that most of what you see here is a function of the GPU on the SoC, not the number of CPU cores, as four of the five phones have 8 CPU cores. Even so, we decided to disable some cores to show the difference in performance between certain core counts in mobile games. First was the Xiaomi Mi 4i.

Figure 4: Fewer Cores, Effect on Frame Rates



The Xiaomi's performance degraded incrementally. But the maximum penalty of going from 8 cores to 2 cores was less than 20% (6.07 FPS to 5.21 FPS in the Hangar test). We reduced the CPU core count by 75%, yet the performance difference was only 15%, clearly indicating that CPU core count is not a factor in smartphone gaming. We followed this test with the G Flex 2 which has a much more powerful CPU and GPU configuration.

Figure 5: Fewer Cores, Higher Frame Rates

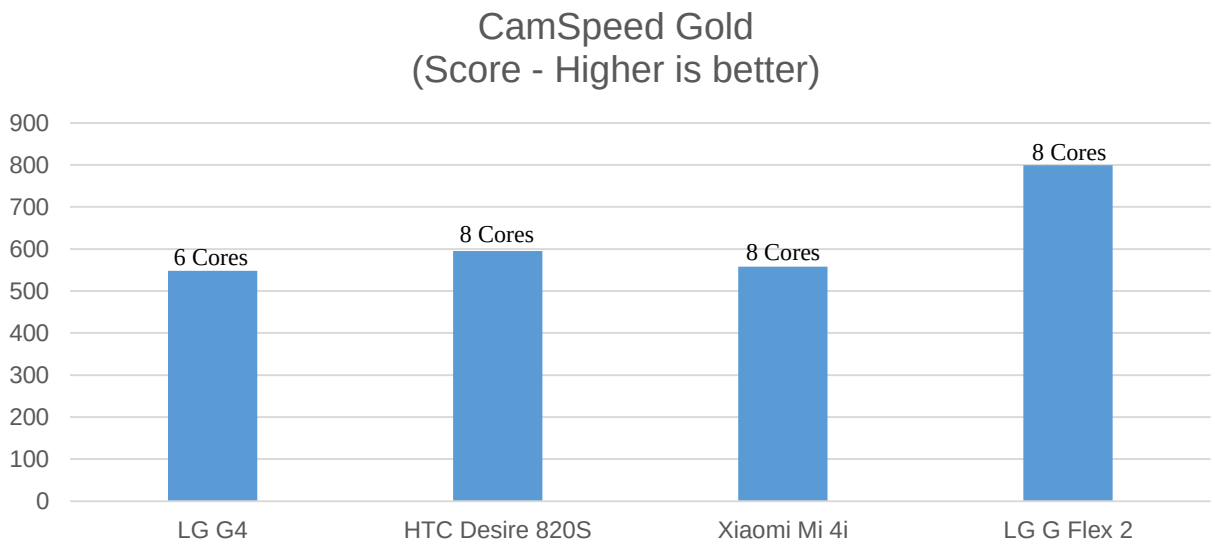


With the LG G Flex 2, **performance improved as we disabled cores**, contrary to the 8 Core Myth. We re-tested it multiple times in Basemark X and either saw no change in the frame rate or a slight increase in frame rate. Realistically, a lot of the variance likely has to do with thermals, although we did wait quite a bit between test runs to make sure there was no thermal throttling. It is possible that having all of the cores active and the GPU active at the same time may cause the GPU to heat up more and cause it to reduce in clock. In our testing with all 8 cores enabled, Basemark X only used between 5 and 6 cores, with performance not really suffering from cores being disabled at all.

CamSpeed Gold

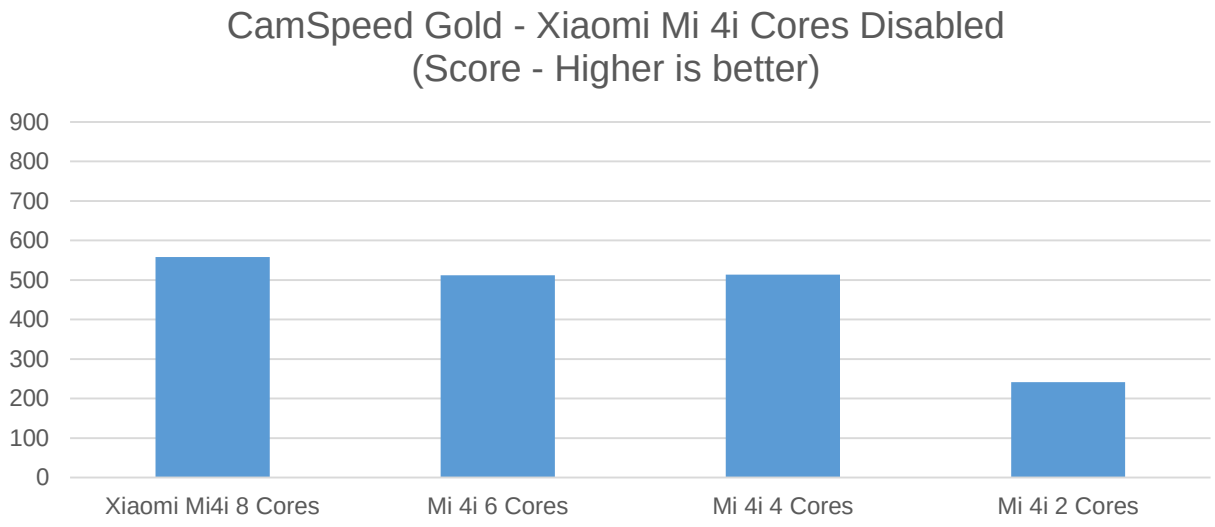
This version of the benchmark is designed for press use and exposes exactly what the times are for each segment of the test. The version we tested with was not compatible with the Samsung Galaxy S6, and as a result we had to strike it from our tests. They have just released a new version that supports the Galaxy S6, but that would require us to retest all of the phones in all of their states. The first batch of testing we did was with the LG G4, HTC Desire 820S, Xiaomi Mi 4i, and LG G Flex 2.

Figure 6: Camera Speeds



Following that, we disabled the cores on the Xiaomi Mi 4i and found **camera performance was mostly unaffected by core count**. With the Mi 4i, like much of our other testing, we went from 8 to 6 to 4 to 2 cores to show scaling. As the chart below indicates, the variance between 8, 6, and 4 cores was less than 10%. However, going down to two cores struck a sharp decrease in overall performance.

Figure 7: Fewer Cores, Effect on Camera Speeds



After those tests, we tested the two applications, WeChat and YouTube.

Application Testing

We tested WeChat and YouTube and disabled cores to see how much the experience would be affected. We found that users could get a playable experience with only 2 cores in both applications, however four cores usually resulted in the smoothest user experience and playback.

The Xiaomi Mi4i, when already in the YouTube application could smoothly play back 1080p content with only 2 cores, however any UI or MIUI calls would cause a slowdown. These issues were not present with 4, 6, or 8 cores, as 4 appears to be the golden number for A53 cores in most scenarios. However, when the Mi4i had 4 cores disabled and 4 cores enabled, it would still kick down to 2 cores. This is because a lot of the video processing is done off CPU on special hardware video decoders on the GPU designed to offload video processing from the CPU. As a result, 2 of ARM's A53s are enough to power 1080p video playback on the Mi4i.

WeChat was less affected by core count than YouTube, since it does not use 1080p video streaming and is mostly messaging and image based. WeChat was very usable with only 2 cores enabled on the Mi4i, however it did seem slower than when 4, 6, or 8 cores were enabled. This doesn't come as much of a surprise considering the overhead of all the other things running on the system. But even so, WeChat is a fairly light application in terms of load and it performs extremely well even with 2 cores.

Call to Action

Through the different benchmarks, tools, and applications, we showed that CPU core count in a modern smartphone is not an accurate measurement of performance or experience. More CPU cores are not always better. We do acknowledge that having many smaller cores is one way to simplify power management, but these tests are not focused on power; they are focused on performance and user experience.

The video and this paper are major steps towards the dispelling the 8 Core Myth and understanding how consumers could be misled about CPU core count, performance, and experience. More testing by press and others will be necessary in order to keep OEMs and SoC vendors honest as hardware and software are constantly evolving.

CPU core counts are not the way that phone manufacturers or carriers should be promoting their devices. CPU core count is only one factor in Android when the SoC has fewer than 4 cores. The marketing of core counts as a primary driver of performance and experience must end and be replaced with improved benchmarking practices and education.

Resources

PCMark for Android v1.1

- PCMark for Android technical guide: http://www.futuremark.com/support/guides
- Futuremark benchmark rules: http://www.futuremark.com/support/benchmark-rules

Basemark X for Android v1.1

- http://www.basemark.com/product-catalog/basemark-x/

CamSpeed for Android v2.31

- http://www.sofica.fi/index.php?page=camspeed-download-gold
- http://www.sofica.fi/CamSpeed2/

Qualcomm Trepn Profiler v6.0

- https://developer.qualcomm.com/software/trepn-power-profiler

3C Toolbox Pro v1.4.1

- http://www.3c71.com/android/?q=node/916
- https://play.google.com/store/apps/details?id=ccc71.at.free

Important Information About This Brief

Inquiries
Please contact us if you would like to discuss this report, and Moor Insights & Strategy will promptly respond.

Citations
This note or paper can be cited by accredited press and analysts but must be cited in-context, displaying author's name, author's title, and "Moor Insights & Strategy". Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

Licensing

Disclosures
Moor Insights & Strategy provides or has provided research, analysis, advising, or consulting to many high-tech companies mentioned in this paper, including ARM, Intel, Nvidia, Qualcomm, and Samsung, cited in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER
The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.