



Reinventing Storage for Social Networks with Intel® Technology

Social network VK engineered a strong storage solution to deliver great performance for its 97 million users and optimize total cost of ownership.



Spotlight on VK

VK is the largest social network in Russia and the Commonwealth of Independent States (CIS) with 97 million monthly active users. Its mission is to connect people, services and companies by creating simple and convenient communication tools. Headquartered in St Petersburg, VK also has bases in Moscow and Sochi, and regional representative programs in Yekaterinburg, Kazan, Nizhny Novgorod, and Kazakhstan.

www.vk.com

Social networking is hugely data intensive. No wonder, then, that data storage consumes a significant part of the budget of VK, Russia's largest social network. VK modernized its tiered storage using Intel® Optane™ persistent memory, Intel® Optane™ SSDs, Intel® SSDs with non-volatile memory express (NVMe), and Intel® FPGA Programmable Acceleration Cards (PACs). As a result, VK expects to make significant financial savings while improving performance.

Challenge

- Lower the cost of data storage, growing at a rate of hundreds of petabytes per year.¹
- Support data tiering, with data migrating to lower cost storage as it is less frequently accessed.
- Eliminate the need to store multiple formats of the same image to serve to different end-user devices.

Solution

- VK re-engineered its storage architecture to lower the cost of storage while meeting its demanding performance requirements.
- VK upgraded the storage for frequently accessed data in its content delivery network (CDN) to Intel® SSDs with 3D NAND technology, and moved the most frequently used data to Intel® Optane™ SSDs.
- VK introduced Intel Optane persistent memory for the rating counter servers that support the newsfeed, migrating data away from more expensive DRAM.
- Intel® field programmable gate arrays (Intel® FPGAs) will be used to convert images on-the-fly from a single high-resolution master copy to the resolution needed for each user – reducing requirement to store multiple image sizes and formats.

Results Reported by VK

- Diverting data from dynamic random-access memory (DRAM) to SSDs and Intel Optane persistent memory running in memory mode significantly cut the cost of storing the hottest data, according to VK.¹
- VK reported that it was able to consolidate servers at a ratio of 2:1 using the new storage solution, supporting the continued data growth, with storage of up to 0.408PB in 1U, reducing power and cooling costs.¹
- Upgrading the processor from the Intel® Xeon® Gold 6230 processor to the Intel® Xeon® Gold 6238R processor cut the compute cost by 40 percent and improved performance per watt by 72 percent¹, according to VK.

At A Glance:

- Optimizing the cost of storage is a business imperative for VK, Russia's leading social network.
- VK engineered a new tiered storage solution. It's based on Intel® Optane™ technology, Intel® 3D NAND SSDs with non-volatile memory express (NVMe), and Intel® FPGA Programmable Acceleration Cards (PACs).
- VK reports it is able to replace two of its old servers with one of its new servers, while improving performance.
- Diverting data from dynamic random-access memory (DRAM) to SSDs and to Intel® Optane™ persistent memory, and resizing images quickly using Intel FGAs, have enabled VK to cut the cost of its hot tier storage.

Cutting the Cost of Storage for Social Networking

Social networking has transformed how we keep in touch with friends, family and colleagues. In Russia and the Commonwealth of Independent States (CIS), the largest social network is VK, which is still growing fast. Friends and families use it to keep in touch, share their daily lives, and celebrate life milestones together. In a typical day, 10 billion messages are exchanged through the platform.

With all the data streaming through the network, it is perhaps no surprise that the data storage infrastructure is one of VK's biggest costs. It accounts for a significant part of the company's annual budget. Optimizing the total cost of ownership (TCO) for storage is a business imperative. VK needs to strike a balance between cost and performance.

Fast storage media are more costly, but deliver a smoother user experience for the most frequently accessed data. In total, 1.1 exabytes of data are distributed across the storage estate. Data is stored close to where it is uploaded. The IT infrastructure behind VK is based on 19,000 servers. There are three main data centers, supported by 30 CDN facilities to speed up access to the hottest data. "Russia is a big country with large distances between cities. We need to have a good CDN cache infrastructure to store data close to users, so that they have a good experience using our social network," said Roman Podpriatov, Deputy COO, VK. "The speed of data access on these servers needs to be very fast."

Original Architecture

VK uses three tiers for caching data on its CDN servers, with data moving down the tiers as it cools. Hot data might be something like a holiday photo that has been recently uploaded to the network and so is still being frequently accessed. Warm data is data that is no longer being accessed so frequently. Typically, it would be up to a month old. Cold data is rarely being accessed.

Before the transition to the new technology, warm data in the CDN was stored on SATA SSDs and hot data was stored in DRAM. Cold data was stored on hard drives in the data center.

Technical Components of Solution

- **Intel® Optane™ SSD DC P4800X.** The Intel Optane SSD DC P4800X enables breakthrough application performance with extremely high throughput, super low latency, and predictably fast service. VK migrated data from DRAM to Intel Optane SSDs on selected CDN servers, reducing the amount of DRAM required.¹ These SSDs are used for the most frequently accessed data to serve it quickly to users, and enable a smoother user experience.
- **Intel® Optane™ persistent memory.** Ratings counters are used for some of the real-time processes that are central to the functioning of the social network, including optimizing the newsfeed. VK uses persistent memory for storage at a lower cost per bit than DRAM.¹
- **Intel® SSD D5-P4320.** These SSDs provide affordable performance for warm data, and play a role in VK's data hierarchy which migrates data from fast to slower (and cheaper) storage, as it is less frequently used. Data in a social network, such as wedding photos, may be hugely popular for a short time and then accessed rarely or never. As data becomes less popular, VK migrates it away from Intel Optane SSDs to more cost-effective Intel SSD D5-P4320, which includes Intel® QLC 3D NAND technology, to lower costs while achieving good performance.
- **Intel® Programmable Acceleration Card (Intel® PAC) with Intel® Arria® 10 GX FPGA.** These PACs reduce storage requirements and provide faster image conversion (size and file type). The CTAccel Image Processor is used to convert the images from a high-resolution master when they are requested by users, avoiding the need for VK to store multiple versions of the same image at different resolutions.
- **Intel® Ethernet Adapter XXV710-DA2.** These 25 Gb/s networking cards provide the strong interconnect bandwidth required for data transfer.

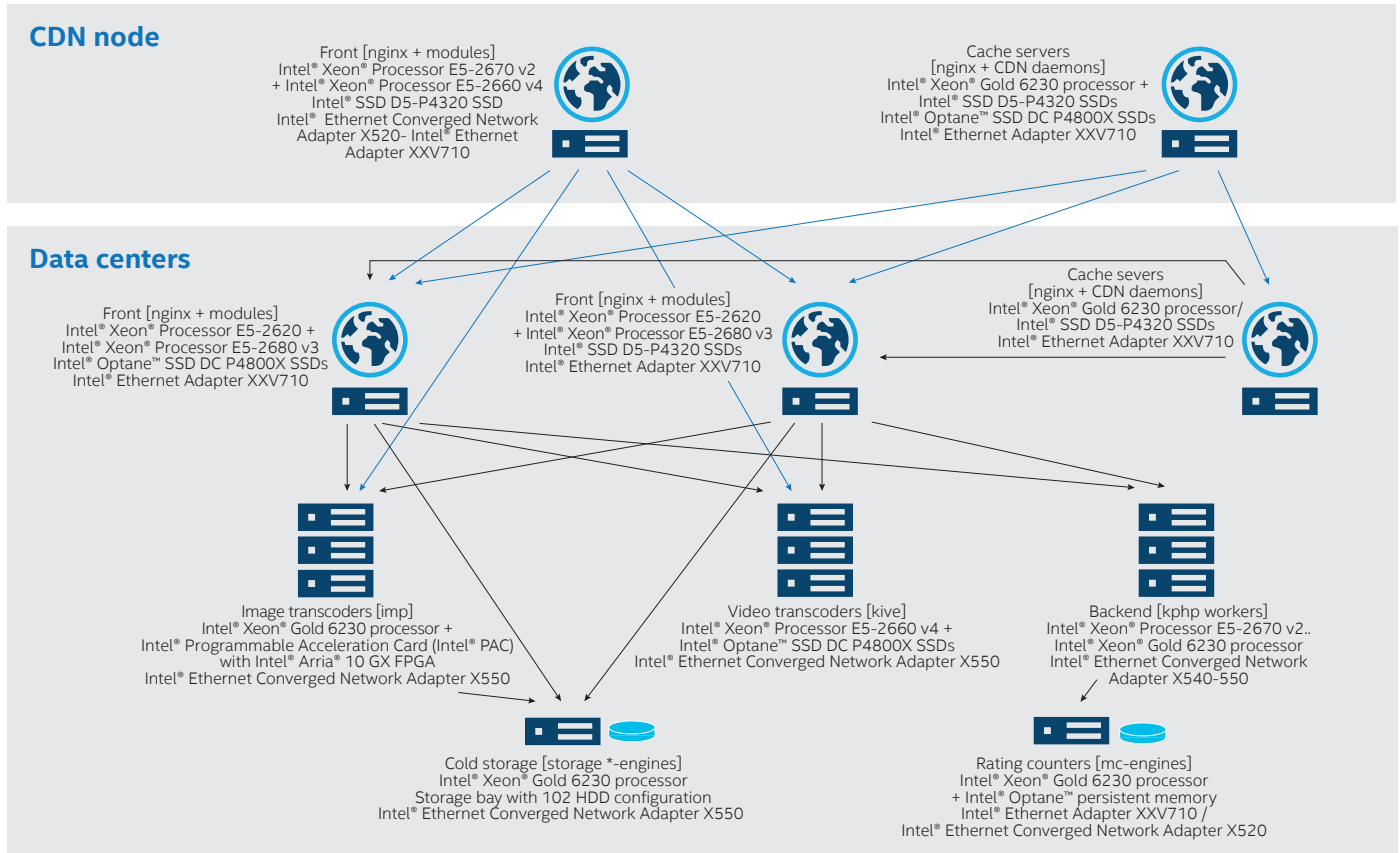


Figure 1. VK's new architecture uses a combination of storage media to meet the performance demands of a modern social network, while lowering the cost of storing infrequently accessed data. This diagram shows the architecture today. Nodes are still being upgraded to the latest processors and networking is being upgraded to 25 Gb/s.

Rating counter servers are used to store and analyze the data, for example, counting the number of likes, rating media, and controlling content demand. On these servers, VK runs a customized version of Memcached. The workloads on these servers did not require any persistent storage (SSDs or hard drives), but were DRAM intensive.

“Our aim was to reduce the number of servers we were using,” said Podpriatov. “If we can reduce the number of rigs we need for our server infrastructure, we can save on our other infrastructure costs too. DRAM was particularly expensive, so we were keen to explore more cost-effective storage options.”

VK was also storing multiple copies of each image to cater for a wide range of user devices, and was keen to explore a more efficient approach.

New Solution Details

VK embarked on a program to modernize its storage architecture. Figure 1 shows the storage architecture today, which continues to evolve as storage, processors and network adapters are upgraded throughout the network.

For CDN cache servers, hot data was moved from expensive DRAM to Intel® Optane™ SSD DC P4800X SSDs, and warm data was moved from SATA SSDs to the Intel® SSD D5-P4320 NVMe drives (see Figure 2).

The CDN cache servers (as shown in Figure 1) are configured with both Intel SSD D5-P4320 drives and Intel Optane SSDs. At the software level, there are containers running separate NGINX instances for each drive type.

NGINX is an open source web server for serving dynamic web content across a network, and VK has optimized its code for better performance under various storage configurations. The “hot data” container based on Intel Optane SSD DC P4800X SSDs is used for caching frequently accessed data, such as new music, new videos or other popular content.

CDN Servers	Warm data	Hot data
Old configuration	SATA SSD	DRAM
New configuration	Intel® SSD D5-P4320	Intel® Optane™ SSD DC P4800X

Figure 2. VK's new storage solution for CDN servers added more performant SSDs for warm data, and lower cost fast storage for hot data.

	2019	2020 forecast	Delta
Processor	Intel® Xeon® processor E5-2660 v4 Intel® Xeon® Gold 6230 processor	Intel® Xeon® Gold 6238R processor	-
Server systems, units	1050	805	(23%)
Cost of computing, \$/queries per second (qps)	4.8	2.9	(40%)
Performance optimization, qps/W	2.9	5	72%

Figure 3. VK’s projected savings using the Intel® Xeon® Gold 6238R processor, compared to the Intel® Xeon® Gold 6230 processor, based on queries per second (qps) per dollar and per Watt, as reported by VK based on VK’s internal performance testing¹. Forecast based on initial engineering analysis and testing, going into production in 2020.

Intel Optane P4800X SSDs help eliminate data center storage bottlenecks and allow bigger, more affordable data sets. The “warm data” container is used for less frequently accessed data, such as images that are about 30 days old, and older video streaming content. CDN servers do not have direct access to cold data on hard drives, which resides in the data center, and is accessed through the data center Front nodes.

“Now we can store both hot and warm data on SSDs, and reduce the amount of DRAM we use,” said Podpriatov. “Previously, our SSDs weren’t fast enough to offer a good user experience for hot data, so we had to keep some data in DRAM. Now, we can put it all on SSDs, which are much cheaper than DRAM.”

In the data center, the front nodes use Intel Optane SSD DC P4800X SSDs for small files, such as avatar images, and Intel SSD D5-P4320 drives for media content. VK is upgrading the processors in these nodes to the Intel Xeon Gold 6238R processor. The software stack is similar to the CDN nodes with the only difference being it’s not containerized and runs on physically different systems to differentiate P4800X and P4320 based front servers. This enables the data center level performance to be delivered.

Intel Optane SSD DC P4800X SSDs are also being used for video transcoding servers. Because there is a very high write load on these servers, SSD endurance is vital. Intel Optane SSDs can provide 60 drive writes per day (DWPD), equivalent to the entire drive being overwritten 60 times per day.

The rating servers are used for counting Likes and ranking newsfeed items, among other things, and these have been upgraded to Intel Optane persistent memory because it offers the performance these real-time processes require while lowering the cost of high-performance storage compared to DRAM. VK is using persistent memory in Memory mode, which does not provide data persistence, but enables VK to benefit from a lower cost per bit than DRAM. The process of upgrading involved testing persistent

memory to see how it performs for VK’s workloads and server configurations, and then deploying it in production. Intel Optane persistent memory uses industry-standard DIMM sockets on the server.

The powerful processors and storage nodes require strong interconnect bandwidth to enable more data to be sent/received than previously, so VK uses two 25 Gb/s Intel® Ethernet Adapter XXV710-DA2 networking cards for each server of disaggregated storage.

The new servers were originally based on the Intel® Xeon® Gold 6230 processor, but VK has since upgraded to the Intel Xeon Gold 6238R processor, which helped VK to increase the performance of storage and compute, optimize TCO and get more performance per watt from the compute capacity.¹

Upgrading the processor cut the compute cost by 40 percent and improved performance per watt by 72 percent (see Figure 3), based on VK’s 2020 forecast. “We saw a significant performance boost when we upgraded,” said Podpriatov. VK is now upgrading the older processors throughout its storage architecture to the Intel Xeon Gold 6238R processor, prioritizing old CPUs with a high core count and high clock speed.

Meanwhile, upgrading from the two 10 Gb/s Ethernet cards that were used previously has enabled a 2.5x extension of data throughput¹, helping to improve the process/store/move balance for each new compute node.

Storage optimization and the need to efficiently process image transcoding algorithms continues to be a growing challenge for VK. To further optimize its storage and increase power efficiency, VK is deploying the Intel® Programmable Acceleration Card (Intel® PAC) with Intel® Arria® 10 GX FPGA (see Figure 4) and running the CTAccel Image Processor workload. The low-power, single slot, low-profile PCIe Intel PAC makes it easy to deploy multiple FPGAs in various VK servers.

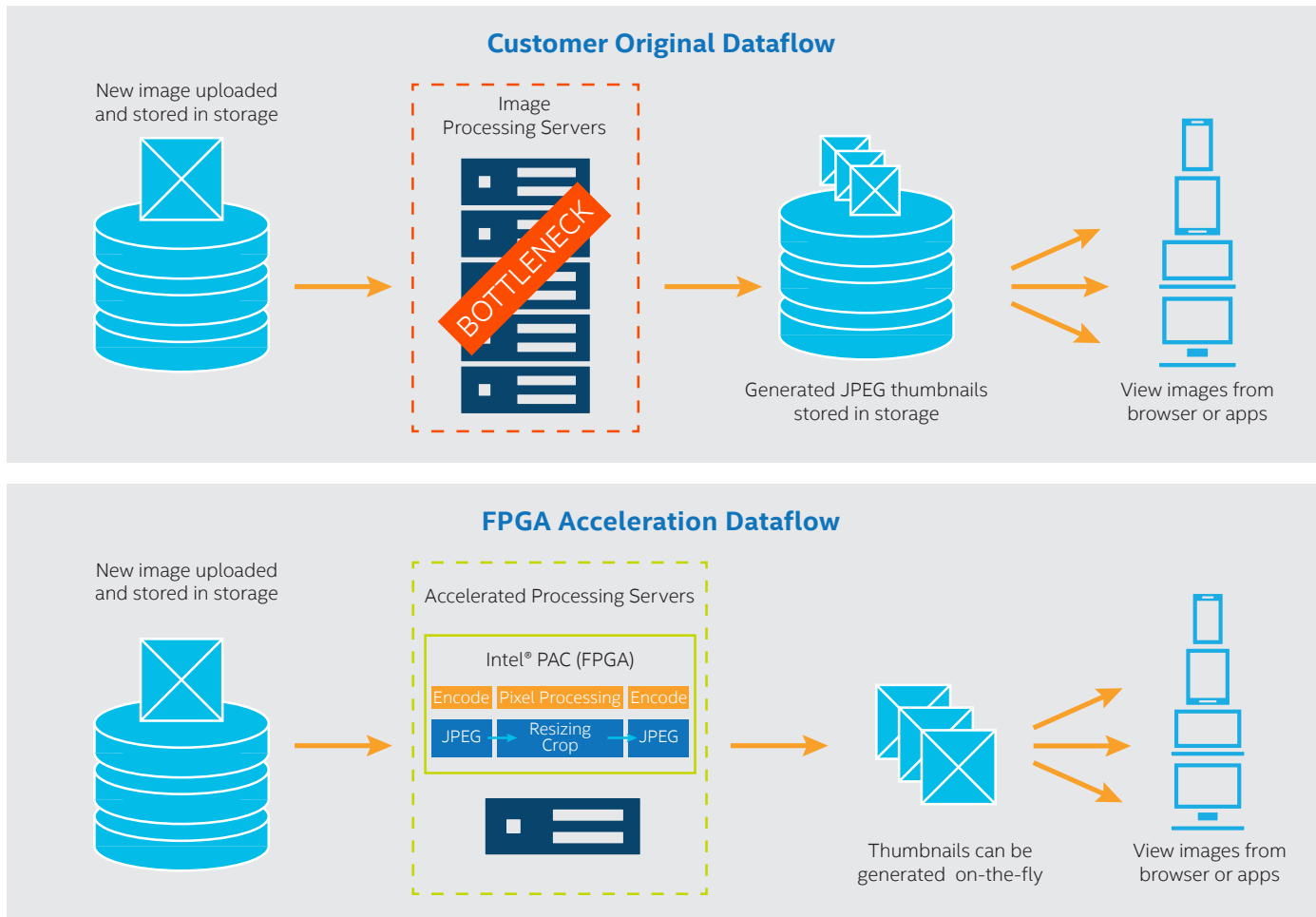


Figure 4. VK dataflow solution with and without using Intel® Programmable Acceleration Card (Intel® PAC). The top image illustrates the need for multiple servers to perform image processing algorithms, and the need for storage post-processing. The bottom image illustrates increased efficiency with workload functions offloaded to the field programmable gate array (FPGA), providing the ability to generate images on-the-fly, reducing storage requirements.

FPGAs can provide custom hardware to accelerate application functions significantly faster than software running on a general-purpose processor. The FPGA is used by VK to convert high-resolution images to the desired size and format on-the-fly. This low-latency, high-throughput solution reduces the overall storage requirements because only high-resolution images need to be stored, instead of multiple copies of the image at different resolutions, and the FPGA improves power efficiency compared to other solutions that VK tested.

Intel Works Closely with Cloud Service Providers

VK and Intel have been working closely together for five years. “During this time, we worked on a lot of projects and fixed a lot of issues together,” said Podpriatov. “We have a good relationship between our companies, and, at VK, we know we can call on Intel if we experience any difficulties during our testing or implementation processes.”

VK executes the implementation itself, but Intel helped with some of the validation processes. “We can only test storage solutions like this in production,” said Podpriatov. “It could take two months to populate an SSD with real data, and to check how data moves from hot to cold storage. It’s impossible to test this in lab conditions.”

“The Intel team helps us all the way from the beginning of a new product to implementation and production,” said Podpriatov. “Intel shares their roadmap and new technologies with us, and we get an opportunity to implement new technologies in our production environment. It gives us a chance to understand whether they’re right for us, and what savings we might achieve as a result of implementing them.”

Business Results Reported by VK

VK estimated it would achieve significant savings when the new storage solution was introduced. There will be ongoing savings in space, power and cooling costs because VK needs fewer racks to store the same volume of data now, with storage of up to 0.4PB in 1U.¹ As a result, VK reports it has reduced its power and cooling costs. “We can replace two of our old servers with one of our new servers, while improving our performance,” said Podpriatov.

Diverting data from DRAM to SSDs and to Intel Optane persistent memory, and resizing images quickly using Intel FPGAs, have enabled VK to cut the cost of its hot tier storage, while allowing Intel to deliver the performance users need efficiently. “We have more performance now at a lower cost than our previous storage solution,” said Podpriatov.

VK in numbers

In 2019, VK had 97 million monthly active users². Every day, users view 9 billion posts and 650 million videos, and they exchange 10 billion messages.¹ They tap the like button a billion times a day.² Over the course of a year, users upload hundreds of petabytes of new data, including photos and videos.¹

Learn More

- [2nd generation Intel® Xeon® Scalable Processors brief](#)
- [Intel® Solid State Drives \(Intel® SSDs\)](#)
- [Intel® Optane™ SSD Series](#)
- [Intel® Optane™ persistent memory](#)
- [Intel® Programmable Acceleration Card \(PAC\) Family](#)
- [Intel® Programmable Acceleration Card with Arria® 10 GX](#)
- [Intel Resources for Cloud Service Providers](#)



Find the solution that is right for your organization.
Contact your Intel representative or visit intel.com/csp.

¹ These results were reported to Intel by VK based on configurations that included the listed Intel components. Tests carried out by VK April-November 2019.

Configurations:

OLD CDN servers: 2 x Intel® Xeon® processor E5-2670 v2 or 2 x Intel® Xeon® processor E5-2680 v4, SATA SSDs, DRAM, 2 x 10 Gb/s Ethernet cards

NEW CDN servers: 2 x Intel® Xeon® Gold 6238R processors or 2 x Intel® Xeon® Gold 6230, Intel® Optane™ SSD DC P4800X SSDs, 6 x Intel® SSD D5-P4320, DRAM, SATA SSD, 2 x 25 Gb/s Intel® Ethernet Adapter XXV710-DA2. Software: cache_api, nginx, Automated Certificate Management Environment (ACME).

OLD rating counters servers: 2 x Intel® Xeon® processor E5-2680 v4, SATA SSD (boot device), DRAM, 2 x 10 Gb/s Ethernet cards

NEW rating counters servers: 2 x Intel® Xeon® Gold 6238R processors or 2 x Intel® Xeon® Gold 6230, 12 x Intel® Optane™ persistent memory, DRAM, SATA SSDs (boot device), 2 x 25 Gb/s Intel® Ethernet Adapter XXV710-DA2. Software: customized version of Memcached.

Microcode for Intel® Xeon® Gold 6230: 0x500002c

Microcode for Intel® Xeon® Gold 6238R: 0x0500002f

² Data from <https://vk.com/about>

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No product or component can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

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 complete information visit www.intel.com/benchmarks

Performance results may not reflect all publicly available security updates. No product or component can be absolutely secure.

Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice Revision #20110804.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Your costs and results may vary.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.