

Sub-linear English Phrase-Checking

This is an example of how (a single shortcut/icon invocation) 'Graphein' r.3 (utilizing 'GAMERA' corpus r.15 x-gram corpora) phrase-checks two incoming .txt files derived from two .htm files published on TOP500 Supercomputing Sites (<http://www.top500.org>).

E:_Gamera_Graphein>dir _TXT-TREE/s

```
04/21/2012 02:56 AM      545,957 About Ivy Bridges and Tegra Hitches.pdf
04/21/2012 03:12 AM        7,916 About Ivy Bridges and Tegra Hitches.pdf.txt
04/21/2012 03:09 AM     212,280 Japan's K Computer.pdf
04/21/2012 03:14 AM      6,683 Japan's K Computer.pdf.txt
```

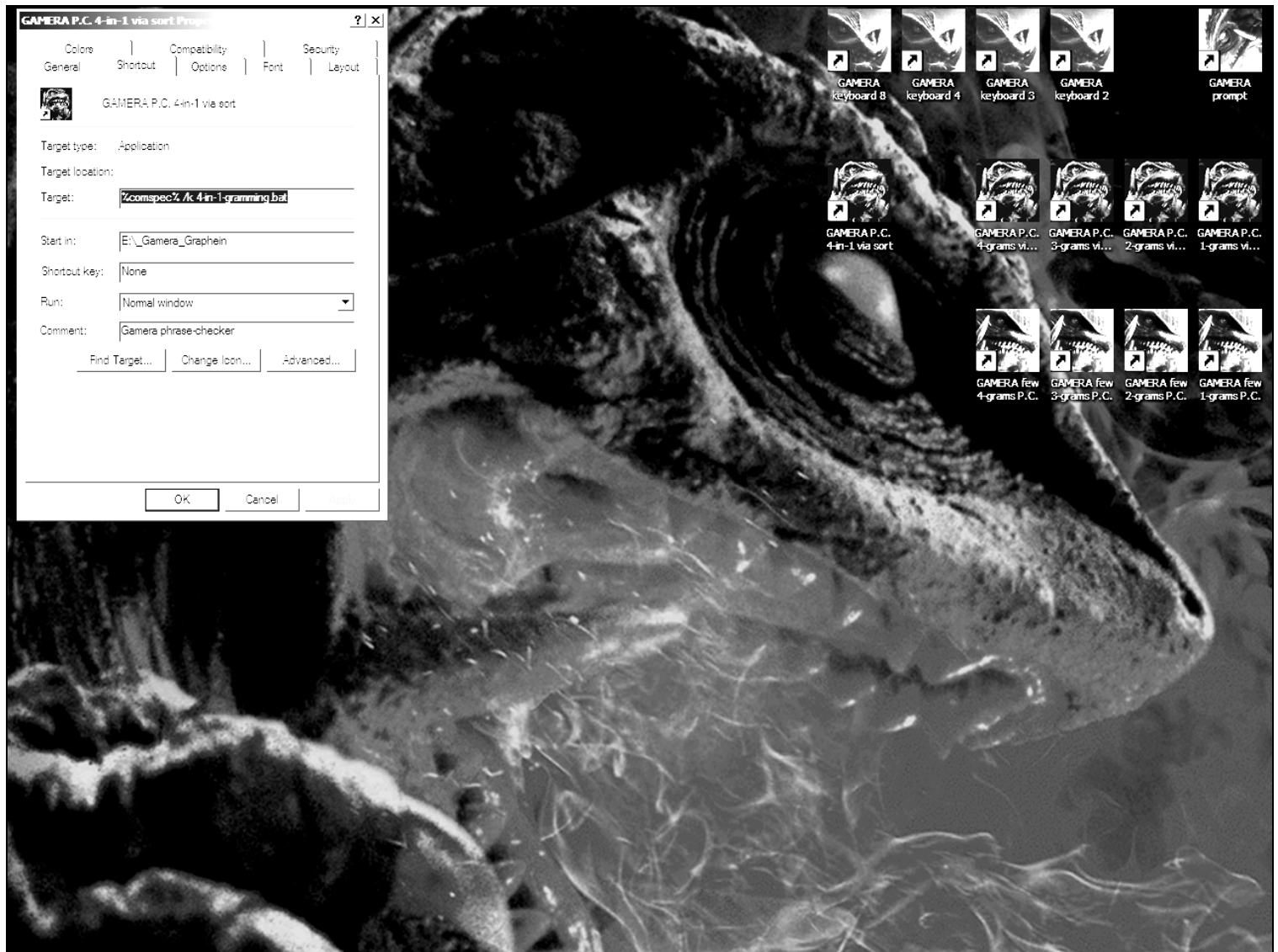
The table holds 'GAMERA' corpus r.15 x-gram corpora stats:

x-grams Order	x-grams Total Number	x-grams Distinct Number	x-grams Distinct Size
1-grams	4,589,933,215	9,181,275	207,307,606 bytes
2-grams	3,835,376,293	124,669,942	3,267,897,913 bytes
3-grams	3,286,567,380	477,829,381	14,630,478,827 bytes
4-grams	2,812,845,037	879,557,846	30,931,053,072 bytes

Here 'GAMERA P.C. 4-in-1 via sort' shortcut having been executed it gives us 16 resultant .TXT files of which 8 are most needed, see below.

The whole task took 4:52:14.78 - 4:08:31.46 = $(4*3600+52*60+14)-(4*3600+8*60+31) = 2623$ seconds.

Note: In next revision of 'Graphein' this time will be reduced significantly (combining Kazuya r19 and Overlapper-Blender_X-GRAM r1+R into one tool thus avoiding unnecessary writes, after that going even further by replacing Kazuya r19 with Kazuya_PTHREADED r0+++_08-threads).



E:_Gamera_Graphein>type TIME.LOG
The current time is: 4:08:31.46

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: 9,181,275 1-grams; 124,669,942 2-grams; 477,829,381 3-grams; 879,557,846 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

at the Oak Ridge National Laboratory, with a 1.75 Petaflop/s performance running the standard Linpack benchmark application. Other top U.S. systems include Cielo, a Cray XE6 at Los Alamos National Laboratory (No. 6); Pleiades, an SGI Altix machine at NASA's Ames Research Center (No.7); Hopper, a Cray XE6 at the National Energy Research Scientific Computing Center (No. 8); and Roadrunner, an IBM system that was the first ever to break the Petaflop/s barrier, at Los Alamos (No. 10). Systems in China, Japan and France round out the Top 10.

Although the top rankings did not change, the newest list does highlight a number of other developments. For example:

- * China keeps increasing its number of systems to 75 and is now clearly the No. 2 country, as a user of HPC, ahead of Japan, UK, France, and Germany.
- * The two Chinese systems at No.2 and No. 4 and the Japanese Tsunami 2.0 system at No. 5 are all using NVIDIA GPUs to accelerate computation.
- * Thirty-nine systems use GPUs as accelerators (up from 17 six months ago), 35 of these use NVIDIA chips, two use Cell processors, and two use ATI Radeon.
- * Already, 62 percent of the systems use processors with six or more cores.
- * Intel continues to provide the processors for the largest share (76.8 percent) of TOP500 systems.
- * Thanks to the K computer Fujitsu captured the No. 2 spot in market share by total performance slightly ahead of Cray, but IBM stays well ahead of both.

With every list, the entry level of performance just to claim the 500th spot increases. In the latest list, the level to the list moved up to the 50.9 Teraflop/s mark on the Linpack benchmark, compared to 39.1 Teraflop/s six months ago. The last system on the newest list was listed at position 305 in the previous TOP500 just six months ago. Total combined performance of all 500 systems has grown to 74.2 Petaflop/s, compared to 58.7 Petaflop/s six months ago and 43.7 Petaflop/s one year ago.

Other points of interest include:

- * A total of 384 systems (76.8 percent) are now using Intel processors. This is slightly down from six months ago (386 systems 77.2 percent).
- * They are now followed by the AMD Opteron family with 63 systems (12.6 percent), down from 66.
- * The share of IBM Power processors has stabilized for now with 49 systems (9.8 percent), up from 45.



About Ivy Bridges and Tegra Hitches.pdf.txt:

Published on /TOP500 Supercomputing Sites/ (<http://www.top500.org>)
[<http://www.top500.org/>](http://www.top500.org/)

About Ivy Bridges and Tegra Hitches

By /admin/
 Created /2012-04-10 03:13/

Andreas Stiller <http://www.top500.org/contributors#stiller> [1]

- * Gigabit Ethernet is still the most-used internal system interconnect technology (223 systems, down from 230 systems), due to its widespread use at industrial customers, followed by InfiniBand technology with 213 systems, up from 208 systems. However, InfiniBand-based systems account for almost twice as much performance (28.7 Petaflop/s) than Gigabit Ethernet ones (14.2 Petaflop/s).
- * IBM kept its lead in systems and has now 223 systems (44.6 percent) compared to HP with 140 systems (28.0 percent). HP had 146 systems (29.2 percent) six months ago, compared to IBM with 218 systems (43.6 percent).

Power consumption of supercomputers

TOP500 now tracks actual power consumption of supercomputers in a consistent fashion.

- * 29 systems on the list are confirmed to use more than 1 megawatt (MW) of power.
- * The No. 1 system, the K computer also reports the highest total power consumption of 12.66 MW. Yet due to its performance, the system is one of the most efficient systems on the list, delivering 830 Mflops/watt.
- * The most energy efficient supercomputers are BlueGene/Q with 2,029 Mflops/watt.
- * Average power efficiency is 282 Mflops/watt (up from 248 Mflops/watt six months ago and 219 Mflops/watt one year ago).
- * Average power consumption of a TOP10 system is 4.56 MW (up from 4.3 MW six months ago) and average power efficiency is 464 Mflops/watt (unchanged).

About the TOP500 List

The TOP500 list is compiled by Hans Meuer of the University of Mannheim, Germany; Erich Strohmaier and Horst Simon of NERSC/Lawrence Berkeley National Laboratory; and Jack Dongarra of the University of Tennessee, Knoxville. For more information, visit www.top500.org
[<http://www.top500.org/>](http://www.top500.org/) [1]

Copyright (c) 2000-2011 TOP500.Org | All trademarks and copyrights on this page are owned by their respective owners.

Source URL: <http://www.top500.org/lists/2011/11/press-release>

Links:
 [1] <http://www.top500.org>

(Translation of the German original in c't by Marcel Sieslack)

Intel causes confusion with the Ivy Bridge's launch date, HP starts a second restructuring attempt and Nvidia launches rumors about a Tegra 4 - all (amongst other things) because of Apple.

While everything is sprouting and leafing, Intel's plants need some extra time this year. And so, a certain ivy-overgrown bridge in the southern English town of Ivybridge still hasn't made it into the spotlight. As for the Ivy Bridge chip's launch date, even Intel has lost track by now: According to an internal diagram with more than half a dozen more or less reasonably scheduled dates, vendors are supposed to

start selling the dual-core processors of the Ivy Bridge generation before they are allowed to talk about them. We are eagerly awaiting the silent movie interludes behind the shop counters. In any case, Intel already gave us something to smile about when it tried to slip us a non-disclosure agreement for Z77 mainboards, which had already arrived in shops.

Apparently, it's not the brand new 22 nm production process with three dimensional FinFET transistors that is to blame for the delays, but the whims of big Intel clients. Allegedly, some of them have made changes to their orders in favor of the low and ultra-low voltage variants (LV and ULV) at the last minute. But only the best of the best silicon dice receive these energy efficiency classifications, and thus it takes longer to produce enough of them. Additionally, it appears that Apple has changed its mind on short notice and now wants chips with HD 4000 instead of HD 2500 graphics, which have 16 instead of only 6 shader units. However, taking into account that Intel is heavily advertising new ultrabooks with Ivy Bridge processors, it would seem a fatal mistake to make the ultrabook manufacturers and their clients wait in favor of Apple.

Meanwhile, the overall Ivy Bridge delay should make the large notebook manufacturers happy: they have a few more weeks to sell off their Sandy Bridge stocks at reasonable prices. The losers are the smaller competitors with empty warehouses. They will now have to equip brand new barebones with old Sandy Bridge processors for some weeks, which will most likely, to put it mildly, confuse the one or other client.

Down the valley

Ultimately, it might also stem from the "damper" the IT market as a whole received because of the hard disk shortage caused by the flood in Thailand: In the fourth quarter, the sales went down by 5.9 percent, which resulted in a meager annual plus of 1.3 percent. Unperturbed by this, Globalfoundries can rejoice over a quarter of a million wafers with 32 nm structures that the former AMD manufacturer has produced by now. And yet, Intel probably took note of this with a quiet smile and is already working on 22 nm Atoms with a powerful graphics core - codename "Valleyview". The GPU is supposed to originate from its large brother, Ivy Bridge, and be four times as fast as the current version. This doesn't mean a lot, but it could solve a pesky problem: Intel doesn't have any 64-bit drivers for the PowerVR graphics of the current Atoms, and the ones for 32 bits are nothing to be happy about.

Angry Like a wolf

Speaking of PowerVR: According to Jon Peddie Research, it's Imagination Technologies, maker of PowerVR, that dominates the SoC market. Reportedly, no other company sells more GPU designs: With a market share of 50 percent the company from the small town of Kings Langley near London easily leaves its competitors Qualcomm (33 percent), Vivante (4.8 percent) and Nvidia (3.2 percent) behind. Jen-Hsun Huang's "super chip" Tegra can currently be found in "superphones" and tablets, but things are not so "super" as far as the market share is concerned. Actually, regarding some topics, nerves seem to be raw in Santa Clara: For instance, Nvidia tried to completely cut us off from all communication concerning the new Kepler graphics card GeForce GTX 680 and also refrained from sending us a sample, because a previously published message simply didn't suit them.

It must have annoyed Nvidia's charismatic boss even more that Apple publicly and offhandedly commented on the weaker graphics performance of the Tegra 3 in comparison to the current iPad PowerVR core. That was at the beginning of March. And now, what a coincidence, there are rumors about the Tegra 4: It could feature a strongly downgraded Kepler architecture, which will have its debut in the GK104 of the desktop graphics card GeForce GTX 680. Supposedly, the mobile ULP GeForce core comes with somewhere between 32 and 64 shaders. Well, the numbers might be right, but why should Nvidia equip Tegra SoCs with cores compatible with shader model 5.1 when version 3.0, scheduled for this summer, of the OpenGL ES interface - the standard for mobile graphics - will only just introduce optional geometry shaders and is mainly derived from the old OpenGL 3? And for the much-discussed CUDA operations, which the Tegra 4 GPU should be able to manage, the unified shaders of model 4.0 are sufficient. For smartphone chips, the mantra is: Leave out what you don't need, because it saves space, reduces costs and makes the battery last longer.

This surely doesn't matter so much with the monster processor Power7+ from IBM. According to the first available pictures, it will pack four CPU dice on a giant ceramic module. It's still unclear, however, what's going to occupy the remaining space. Newcomer Adapteva would surely say: Cores, cores and more cores. Its processor Epiphany E4K has 4096 cores - admittedly, tiny ones - and is supposed to reach 5.632 teraflops at 700 MHz of clock speed while, nonetheless, consuming only 80 watts. If correct, the chip, produced in 28 nm structures by Globalfoundries, would be more efficient than all current GPGPU implementations or competitors like Larrabee and Knights Corner.

HP is meanwhile dreaming of equipping its own workstations with that knightly power. While the head of HP's department for workstation and thin clients, Jim Zarafana - obviously proud - presented new models and merrily philosophized about possible future workstation tablets with us, his boss started restructuring the company, again. Zarafana won't be affected though, as his business sector is growing quickly - even if mainly at the expense of the competition. However, the HP dinosaur "VJ" (Vyomesh Joshi) is being consigned to history after 31 years of service so that his Printing Group can be merged with the PC branch under the lead of Todd Bradley. After the failed attempt with Touchpad and WebOS, this branch of the company would be well advised to quickly pull an iPad competitor out of its hat. Allegedly, HP - maybe with a partner -- is working on a Windows 8 tablet.

According to rumors, the upcoming Tegra 4 combined processors will feature a graphics unit with Kepler architecture. Nvidia only just introduced this architecture for graphics cards at the end of March, with the GK104 chip (see picture).

* whispers <<http://www.top500.org/tags/whispers>>

Copyright (c) 2000-2011 TOP500.Org | All trademarks and copyrights on this page are owned by their respective owners.

*Source URL: *
http://www.top500.org/blog/2012/04/10/about_ivy_bridges_and_tegra_hitches

*Links: *
[1] <http://www.top500.org/contributors#stiller>



The first half of next phrase-checked (with **Gamera** corpus) 1-gram wordlist contains unfamiliar 1-grams whereas the second half holds all familiar to **Gamera** 1-grams with their corresponding occurrences:

Overlapped lines, i.e. lines common for both files: **801**

Unfamiliar lines, i.e. lines from 1st file not encountered in 2nd file: **5**

adapteva
sieslack
ultrabook
ultrabooks
zarafana

a	9,999,999	be	9,999,999	coincidence	0,027,478
able	1,029,006	because	3,774,317	combined	0,175,814
about	5,737,558	before	4,191,612	comes	0,716,597
accelerate	0,011,165	began	1,038,325	commented	0,028,617
accelerators	0,003,281	beginning	0,597,365	communication	0,284,523
according	0,819,951	behind	0,864,337	company	1,077,147
account	0,924,676	being	3,352,434	compared	0,276,366
achieve	0,139,600	benchmark	0,014,004	comparison	0,208,618
achieved	0,139,284	berkeley	0,044,094	compatible	0,061,205
actual	0,307,971	best	1,305,612	competition	0,080,386
actually	0,444,837	between	2,752,385	competitor	0,013,415
additionally	0,037,559	big	0,594,619	competitors	0,025,910
admin	0,047,424	birds	0,230,675	compiled	0,075,570
admittedly	0,007,111	bit	0,573,991	completely	0,329,643
advanced	0,376,902	bits	0,153,020	computation	0,035,177
advertising	0,366,096	blame	0,118,888	computational	0,027,351
advised	0,074,812	blog	0,018,380	computer	0,677,316
affected	0,236,455	bluegene	0,000,014	computing	0,085,782
after	4,649,077	boss	0,056,006	concerned	0,247,125
again	2,741,213	both	2,510,656	concerning	0,314,209
ago	0,434,285	bradley	0,019,325	conference	0,137,777
agreement	0,166,149	branch	0,193,260	confirmed	0,120,395
ahead	0,211,603	brand	0,067,741	confuse	0,018,260
aics	0,000,196	break	0,377,461	confusion	0,159,616
alamos	0,001,724	bridge	0,233,251	consigned	0,009,490
all	9,999,999	bridges	0,044,819	consistent	0,123,951
allegedly	0,006,201	brother	0,644,433	consuming	0,034,763
allowed	0,408,565	build	0,301,083	consumption	0,081,767
almost	1,222,049	built	0,435,123	continues	0,117,830
already	1,110,147	business	1,214,788	contrary	0,259,260
also	4,931,597	but	9,999,999	contributors	0,015,822
although	1,170,336	by	9,999,999	copyright	0,754,171
altix	0,000,008	c	4,959,448	copyrights	0,004,540
amd	0,015,502	calculations	0,058,961	core	0,191,616
ames	0,009,494	calif	0,011,192	cores	0,007,937
amongst	0,147,541	called	2,138,308	corner	0,288,393
an	9,999,999	can	9,999,999	correct	0,281,617
and	9,999,999	captured	0,090,378	costs	0,189,430
andreas	0,010,784	card	0,241,950	could	5,278,103
angry	0,218,378	cards	0,144,446	counters	0,023,971
annoyed	0,035,972	case	1,636,277	country	1,186,860
annual	0,140,911	caused	0,395,394	cpu	0,081,673
any	5,915,919	causes	0,371,910	cray	0,003,346
apparently	0,220,673	cell	0,702,221	created	0,647,998
appears	0,529,970	center	0,535,145	cuda	0,000,077
apple	0,094,281	ceramic	0,007,419	current	0,682,894
application	1,100,985	certain	1,216,767	currently	0,164,487
architecture	0,201,885	change	1,251,445	customers	0,163,844
are	9,999,999	changed	0,421,315	cut	0,581,421
arrived	0,309,596	changes	0,668,784	damper	0,003,246
as	9,999,999	charismatic	0,005,066	data	2,403,299
at	9,999,999	china	0,178,027	date	0,701,708
ati	0,009,045	chinese	0,168,753	dates	0,108,659
atoms	0,062,560	chip	0,045,090	debut	0,009,492
atop	0,013,366	chips	0,031,665	delay	0,170,993
attempt	0,352,547	cielo	0,005,881	delays	0,028,076
available	0,783,347	claim	0,233,017	delivering	0,031,313
average	0,255,486	clara	0,048,284	department	0,266,947
awaiting	0,040,043	classifications	0,012,148	derived	0,238,576
back	2,992,197	clearly	0,344,578	designs	0,079,582
barebones	0,000,464	client	0,575,215	desktop	0,118,980
barrier	0,072,712	clients	0,174,736	dethroned	0,003,500
based	0,960,726	clock	0,317,486	developments	0,044,731
battery	0,076,729	codename	0,000,497	diagram	0,082,927

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked: Downloaded from www.sanmayce.com

dice	0,028,460	found	2,340,740	information	1,934,682
did	3,982,822	four	1,245,635	installed	0,206,262
didn	0,754,752	fourth	0,539,405	instance	0,559,340
dimensional	0,073,123	france	0,452,160	instead	0,684,246
dinosaur	0,004,420	from	9,999,999	institute	0,129,316
disclosure	0,021,953	fujitsu	0,001,461	intel	0,045,229
discussed	0,245,415	full	1,308,772	interconnect	0,007,428
discussion	0,255,105	future	0,634,745	interest	0,673,865
disk	0,379,019	gave	1,118,152	interface	0,496,346
does	2,153,468	geforce	0,002,150	interludes	0,001,754
doesn	0,402,041	generation	0,232,719	internal	0,337,395
dominates	0,007,887	geometry	0,052,664	into	7,018,941
don	2,713,719	german	0,357,820	introduce	0,084,074
dongarra	0,000,189	germany	0,214,552	introduced	0,228,763
down	3,458,012	giant	0,114,328	ipad	0,000,228
downgraded	0,000,706	gigabit	0,006,824	is	9,999,999
dozen	0,158,393	gk	0,008,335	it	9,999,999
dreaming	0,058,558	globalfoundries	0,000,002	its	5,944,195
drivers	0,079,951	going	1,644,338	ivy	0,020,126
dual	0,057,515	gpgpu	0,000,009	ivybridge	0,000,036
due	0,605,735	gpu	0,001,623	jack	0,233,102
e	5,574,737	gpus	0,000,409	jaguar	0,004,432
eagerly	0,082,696	graphics	0,108,134	japan	0,115,149
easily	0,470,949	group	1,138,581	japanese	0,131,503
edition	0,956,797	growing	0,262,374	jen	0,006,937
editor	0,225,646	grown	0,201,568	jim	0,118,160
efficiency	0,084,267	gtx	0,000,984	jon	0,020,682
efficient	0,116,860	had	9,999,999	joshi	0,001,801
empty	0,327,435	half	1,254,173	june	0,227,019
end	2,032,054	hans	0,061,091	just	3,359,395
energy	0,530,375	happy	0,520,360	k	1,673,683
english	2,145,609	hard	0,837,970	keeps	0,112,916
enough	1,350,514	has	8,056,845	kepler	0,006,069
entry	0,349,812	hasn	0,057,441	kept	0,633,432
epiphany	0,004,391	hat	0,307,894	kings	0,236,495
equip	0,009,674	have	9,999,999	knightly	0,007,079
equipping	0,002,040	hd	0,024,687	knights	0,078,801
erich	0,003,422	head	1,877,212	knoxville	0,003,117
es	0,471,778	heavily	0,106,630	kobe	0,001,439
ethernet	0,071,786	held	0,800,370	laboratory	0,131,776
even	3,816,045	held	0,800,370	langley	0,005,109
ever	1,786,564	highest	0,302,958	large	1,330,908
every	2,695,854	highlight	0,030,852	largest	0,129,929
everything	0,818,585	his	9,999,999	larrabee	0,000,397
example	2,025,822	history	2,058,566	last	2,240,801
expense	0,155,436	hitches	0,000,804	latest	0,098,640
extra	0,143,529	hopper	0,007,147	launch	0,065,875
fact	1,184,488	horst	0,001,931	launches	0,009,085
failed	0,264,060	however	2,157,725	lawrence	0,362,412
family	0,971,353	hp	0,047,833	lead	0,465,849
far	1,676,146	hpc	0,002,458	leafing	0,000,554
fashion	0,227,317	hsun	0,000,296	leave	0,851,142
fast	0,417,173	http	0,705,760	leaves	0,311,921
fatal	0,126,404	huang	0,010,979	less	1,438,038
favor	0,160,214	ibm	0,283,383	level	0,909,178
feather	0,037,094	if	9,999,999	like	5,189,218
feature	0,270,512	imagination	0,178,624	likely	0,497,862
few	1,657,177	implementations	0,041,551	links	0,182,606
finfet	0,000,010	impressive	0,043,536	linpack	0,000,287
first	4,693,871	in	9,999,999	list	1,011,917
flood	0,100,153	include	0,779,332	listed	0,154,096
followed	0,700,763	increases	0,165,300	lists	0,188,008
for	9,999,999	increasing	0,218,808	london	0,590,913
former	0,410,278	industrial	0,117,288	longer	0,711,440
		infiniband	0,000,595		

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked: Downloaded from www.sanmayce.com

los	0,181,611	newcomer	0,010,249	plus	0,353,363
losers	0,005,891	newest	0,012,265	points	0,433,285
lost	0,800,009	next	1,919,186	position	0,760,232
lot	0,410,354	nine	0,266,706	possible	1,212,990
low	0,870,627	nm	0,044,779	power	1,743,760
lv	0,036,957	no	9,999,999	powerful	0,307,442
m	5,067,676	non	0,719,908	powervr	0,000,016
machine	0,390,775	nonetheless	0,032,287	presented	0,296,186
made	3,918,178	not	9,999,999	press	1,118,331
mainboards	0,000,058	note	0,989,875	previous	0,759,874
mainly	0,118,932	notebook	0,020,292	previously	0,204,118
maintained	0,155,820	nothing	2,135,449	prices	0,092,869
make	3,302,629	notice	0,375,788	printing	0,111,050
maker	0,065,953	nov	0,053,772	probably	0,673,629
makes	0,708,223	now	5,646,675	problem	0,704,544
manage	0,207,909	number	2,187,737	process	1,215,611
mannheim	0,003,615	numbers	0,422,255	processing	0,272,007
mantra	0,025,263	nvdiia	0,003,909	processor	0,118,548
manufacturer	0,052,132	oak	0,069,084	processors	0,061,202
manufacturers	0,043,114	obviously	0,124,228	produce	0,408,666
many	3,655,034	occupy	0,058,414	produced	0,411,412
marcel	0,011,088	of	9,999,999	production	0,354,235
march	0,354,345	off	2,081,967	proud	0,188,355
mark	0,519,928	offhandedly	0,000,411	provide	0,618,298
market	0,339,141	old	2,893,790	publicly	0,051,678
matter	1,130,975	on	9,999,999	published	0,548,855
maybe	0,256,588	one	9,999,999	publishing	0,126,062
meager	0,006,935	ones	0,366,190	pull	0,146,914
mean	0,891,398	only	6,489,190	put	1,759,759
meanwhile	0,124,119	opengl	0,011,972	q	0,686,604
megawatt	0,000,493	operations	0,307,053	quadrillion	0,000,454
merged	0,027,332	opteron	0,000,689	qualcomm	0,000,843
merrily	0,017,699	optional	0,112,193	quarter	0,187,251
message	0,720,247	or	9,999,999	quickly	0,423,762
meuer	0,000,057	orders	0,314,817	quiet	0,300,502
mflops	0,000,228	org	0,301,045	radeon	0,001,584
mhz	0,034,561	original	0,499,655	rankings	0,002,851
might	2,840,468	originate	0,019,999	raw	0,116,648
mildly	0,022,419	other	6,997,180	reach	0,410,042
million	0,294,134	out	7,047,551	reasonable	0,136,629
mind	2,068,941	over	4,020,615	reasonably	0,045,489
minute	0,293,868	overall	0,141,661	receive	0,500,153
mistake	0,163,712	overgrown	0,010,834	received	0,785,712
mobile	0,132,488	own	3,350,758	recent	0,270,847
model	0,609,170	owned	0,087,687	reduces	0,060,202
models	0,219,247	owners	0,066,054	reflects	0,047,973
module	0,202,209	p	4,629,606	refrained	0,010,487
monster	0,068,581	pack	0,118,169	regarding	0,163,894
months	0,513,296	page	1,621,654	rejoice	0,073,312
more	8,221,818	partially	0,066,885	release	0,338,511
most	4,198,860	partner	0,104,736	remain	0,462,883
moved	0,442,160	pc	0,143,375	remaining	0,179,248
movie	0,109,653	peddie	0,000,144	reportedly	0,008,393
much	3,505,868	per	0,789,784	reports	0,271,617
must	4,187,397	percent	0,318,715	research	0,479,390
mw	0,018,269	performance	0,529,583	respective	0,064,369
nasa	0,018,444	pesky	0,001,581	restructuring	0,008,482
national	0,550,461	petaflop	0,000,013	resulted	0,075,130
near	0,903,503	philosophized	0,000,685	ridge	0,059,409
nearest	0,097,842	picture	0,346,446	right	2,547,895
need	1,775,137	pictures	0,154,682	riken	0,000,514
nersc	0,000,053	place	2,392,540	roadrunner	0,000,396
nerves	0,087,787	plants	0,211,348	round	0,815,363
new	4,474,803	pleiades	0,002,808	rumors	0,017,893

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked: Downloaded from www.sanmayce.com

running	0,580,660	standard	0,626,504	todd	0,020,730
s	9,999,999	start	0,812,578	took	1,534,038
said	8,585,256	started	0,482,729	top	0,633,515
sales	0,234,589	starts	0,149,254	topic	0,102,548
sample	0,237,497	stay	0,390,077	topics	0,100,059
sandy	0,049,596	stays	0,027,386	tops	0,043,571
santa	0,048,671	stem	0,089,420	total	0,527,116
saves	0,033,452	still	2,536,171	touchpad	0,000,562
say	2,877,594	stiller	0,001,867	town	0,626,919
sc	0,082,779	stocks	0,047,285	track	0,202,732
scheduled	0,041,747	strohmaier	0,000,018	tracks	0,070,733
science	0,533,379	strongly	0,154,181	trademarks	0,022,703
scientific	0,249,341	structures	0,216,316	transistors	0,005,458
seattle	0,013,702	sufficient	0,302,182	translation	0,230,633
second	1,477,591	suit	0,153,432	trends	0,050,507
sector	0,071,394	summer	0,281,372	tried	0,497,436
see	5,979,234	super	0,079,424	tsubame	0,000,006
seem	0,585,742	supercomputer	0,001,811	tuesday	0,044,564
sell	0,166,334	supercomputers	0,001,512	turnover	0,018,627
selling	0,083,808	supercomputing	0,001,754	twice	0,239,647
sells	0,018,083	superphones	0,000,001	two	5,106,448
sending	0,178,538	supposed	0,301,371	u	1,741,815
service	1,196,294	supposedly	0,014,661	uk	0,072,420
session	0,274,888	surely	0,268,355	ulp	0,000,373
sgi	0,003,052	system	2,520,444	ultimately	0,099,880
shader	0,007,272	systems	0,803,037	ultra	0,029,750
shaders	0,002,227	t	9,999,999	ulv	0,000,286
share	0,368,091	tablet	0,024,370	unchanged	0,035,563
shop	0,482,294	tablets	0,031,594	unclear	0,027,598
shops	0,047,034	tags	0,094,443	under	2,496,958
short	1,204,266	taken	1,165,686	unified	0,024,089
shortage	0,011,154	takes	0,472,437	unit	0,279,939
should	5,206,380	taking	0,624,356	units	0,177,969
showed	0,366,592	talk	0,613,808	university	0,771,985
silent	0,337,971	technologies	0,114,070	unperturbed	0,002,576
silicon	0,019,920	technology	0,337,068	up	7,038,147
simon	0,103,771	tegra	0,000,028	upcoming	0,007,952
simply	0,686,028	tenn	0,002,947	url	0,170,103
since	1,703,756	tennessee	0,033,631	us	4,203,755
sites	0,271,464	teraflop	0,000,060	use	4,219,545
six	0,677,839	teraflops	0,000,073	used	2,965,756
slightly	0,205,035	th	1,347,849	user	1,320,384
slip	0,089,139	thailand	0,013,928	using	1,959,838
small	1,674,043	than	6,084,569	utilize	0,024,187
smaller	0,217,405	thanks	0,189,561	valley	0,225,579
smartphone	0,003,608	that	9,999,999	valleyview	0,000,014
smile	0,351,225	the	9,999,999	variants	0,029,379
so	9,999,999	their	9,999,999	vendors	0,044,431
soc	0,065,850	them	9,051,881	version	0,653,304
socs	0,000,649	there	9,999,999	very	4,875,767
solve	0,084,740	these	6,333,983	visit	0,338,324
some	6,343,203	they	9,999,999	vivante	0,001,124
something	1,889,003	thin	0,227,423	vj	0,006,516
somewhere	0,186,713	things	2,485,780	voltage	0,064,104
source	0,773,397	thirty	0,320,183	vyomesh	0,000,008
southern	0,192,847	this	9,999,999	wafers	0,003,502
space	0,758,776	though	2,087,593	wait	0,434,232
sparc	0,007,257	three	2,434,599	wants	0,261,306
speaking	0,375,302	thus	1,715,289	warehouses	0,010,740
speed	0,304,548	tianhe	0,000,010	was	9,999,999
spot	0,266,512	time	6,393,502	watt	0,014,208
spotlight	0,005,545	times	1,155,496	watts	0,023,531
sprouting	0,004,585	tiny	0,122,363	we	9,999,999
stabilized	0,010,821	to	9,999,999	weaker	0,035,268

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked: Downloaded from www.sanmayce.com

webo\$	0,000,046	who1e	1,650,166	wor1d	3,481,580
weeks	0,325,114	why	1,976,895	wou1d	8,053,220
we11	4,190,475	widespread	0,055,110	www	1,122,242
went	1,980,138	wi11	9,999,999	xe	0,010,434
what	9,054,758	windows	0,962,633	xt	0,012,375
when	9,999,999	with	9,999,999	year	1,281,682
which	9,999,999	wolf	0,094,601	year1y	0,026,297
whi1e	2,655,601	won	0,635,375	years	2,240,478
whims	0,007,176	working	0,558,836	yet	2,344,763
whispers	0,020,332	workstation	0,038,329	you	9,999,999
who	9,401,695	workstations	0,016,726	z	0,674,446

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: 9,181,275 1-grams; 124,669,942 2-grams; 477,829,381 3-grams; 879,557,846 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

The first half of next phrase-checked (with **Gamera** corpus) 2-gram wordlist contains unfamiliar 2-grams whereas the second half holds all familiar to **Gamera** 2-grams with their corresponding occurrences:

Overlapped lines, i.e. lines common for both files: **1367**

Unfamiliar lines, i.e. lines from 1st file not encountered in 2nd file: **159**

a_tegra
accelerate_computation
adapteva_would
alamos_no
altix_machine
amd_manufacturer
an_ipad
and_tegra
and_webos
andreas_stiller
annoyed_nvidia
annual_plus
apple_publicly
are_bluegene
barebones_with
big_intel
bluegene_q
bridge_delay
bridge_processors
bridge_stocks
by_globalfoundries
ceramic_module
certain_ivy
chinese_tianhe
chip_tegra
codename_valleyview
competitors_qualcomm
computer_tops
contributors_stiller
core_codename
cores_admittedly
cpu_dice
cray_xe
cuda_operations
current_atoms
current_gpgpu
current_ipad
delivering_mflops
dice_receive
dimensional_finfet
dinosaur_vj
discussed_cuda
dongarra_of
downgraded_kepler
editor_erich
efficiency_classifications
efficient_supercomputers
equip_brand
equip_tegra
finfet_transistors
from_mflops
fujitsu_captured
geforce_core
giant_ceramic
gk_chip
globalfoundries_can
gpgpu_implementations
hans_meuer

hitches_by
hitches_links
horst_simon
hp_dinosaur
hp_maybe
in_superphones
include_cielo
intel_causes
interludes_behind
introduce_optional
ipad_competitor
ipad_powervr
ivy_bridges
ivybridge_still
japanese_tsubame
jim_zarafana
kepler_architecture
kepler_graphics
kings_langley
knights_corner
launches_rumors
like_larrabee
manufacturers_happy
marcel_sieslack
meanwhile_dreaming
merrily_philosophized
meuer_of
mflops_watt
million_wafers
mobile_ulp
movie_interludes
mw_six
nersc_lawrence
new_kepler
new_ultrabooks
newcomer_adapteva
nvidia_equip
nvidia_launches
of_petaflop
of_powervr
offhandedly_commented
overall_ivy
overgrown_bridge
pc_branch
peddie_research
petaflop_s
powervr_core
powervr_graphics
processor_epiphany
qualcomm_percent
reasonably_scheduled
restructuring_attempt
riken_advanced
sandy_bridge
science_aics
should_nvidia
sieslack_intel
silicon_dice

smartphone_chips
sparc_processing
started_restructuring
stiller_http
stiller_published
stiller_translation
strongly_downgraded
supercomputers_top
superphones_and
tags_whispers
tegra_all
tegra_can
tegra_combined
tegra_gpu
tegra_hitches
tegra_in
tegra_soc
teraflop_s
teraflops_at
the_tegra
the_tianhe
the_ultrabook
tianhe_a
tops_petaflop
ulp_geforce
ultrabook_manufacturers
ultrabooks_with
ulp_at
upcoming_tegra
valley_ultimately
variants_lv
vj_vyomesh
voltage_variants
watt_six
watt_unchanged
weaker_graphics
whispers_copyright
whispers_http
wolf_speaking
workstation_tablets
z_mainboards
zarafana_obviously
zarafana_won

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: 9,181,275 1-grams; 124,669,942 2-grams; 477,829,381 3-grams; 879,557,846 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

a_birds	0,000,134	all_current	0,002,394	are_confirmed	0,008,640
a_certain	0,394,917	all_systems	0,005,482	are_eagerly	0,000,564
a_coincidence	0,006,166	all_trademarks	0,001,058	are_not	1,065,625
a_consistent	0,015,184	all_using	0,000,534	are_nothing	0,016,099
a_cray	0,000,198	allowed_to	0,161,672	are_now	0,137,585
a_dozen	0,121,173	almost_twice	0,001,117	are_owed	0,002,292
a_fatal	0,013,479	already_arrived	0,001,833	are_rumors	0,000,301
a_feather	0,009,532	already_gave	0,000,236	are_sufficient	0,007,317
a_few	1,140,239	already_working	0,000,816	are_supposed	0,017,526
a_full	0,134,310	also_one	0,008,998	are_the	0,849,233
a_giant	0,021,278	also_refrained	0,000,033	arrived_in	0,034,205
a_graphics	0,003,897	also_reports	0,000,824	as_accelerators	0,000,036
a_lot	0,258,571	also_stem	0,000,087	as_far	0,230,198
a_market	0,015,214	although_the	0,187,779	as_fast	0,042,434
a_meager	0,000,924	amd_opteron	0,000,183	as_for	0,213,706
a_million	0,047,598	ames_research	0,000,292	as_his	0,208,353
a_non	0,055,505	amongst_other	0,004,094	as_its	0,075,281
a_number	0,370,710	an_job	0,006,576	as_much	0,353,469
a_partner	0,012,785	an_impressive	0,006,937	as_powerful	0,004,268
a_performance	0,013,398	an_internal	0,026,962	as_the	2,943,712
a_pesky	0,000,148	an_sgi	0,000,085	at_industrial	0,000,114
a_powerful	0,056,637	and_also	0,190,261	at_los	0,001,991
a_previously	0,008,007	and_average	0,002,932	at_mhz	0,001,979
a_quarter	0,068,813	and_be	0,194,594	at_nasa	0,000,575
a_quiet	0,040,598	and_copyrights	0,000,302	at_no	0,048,229
a_sample	0,037,709	and_for	0,448,362	at_position	0,004,459
a_second	0,261,125	and_france	0,016,251	at_reasonable	0,001,008
a_strongly	0,004,437	and_germany	0,010,256	at_sc	0,000,069
a_system	0,122,361	and_has	0,193,034	at_the	5,075,084
a_top	0,015,301	and_horst	0,000,082	ati_radeon	0,000,402
a_total	0,066,207	and_installed	0,005,433	atoms_with	0,000,712
a_user	0,131,998	and_is	0,578,756	atop_the	0,005,145
a_whole	0,154,156	and_jack	0,010,067	atop_top	0,000,001
a_windows	0,040,899	and_knights	0,002,521	attempt_and	0,001,295
a_wolf	0,013,762	and_leafing	0,000,021	attempt_with	0,000,809
able_to	0,903,497	and_makes	0,041,729	available_pictures	0,000,005
about_a	0,190,221	and_mannheim	0,000,111	average_power	0,000,704
about_ivy	0,000,021	and_merrily	0,000,909	awaiting_the	0,009,314
about_possible	0,001,581	and_mflops	0,000,007	back_in	0,127,940
about_the	1,027,440	and_more	0,393,644	based_systems	0,006,918
about_them	0,073,972	and_no	0,265,509	battery_last	0,000,018
about_when	0,005,905	and_now	0,372,705	be_able	0,420,376
accelerators_up	0,000,001	and_nvidia	0,000,241	be_affected	0,019,040
according_to	0,794,443	and_offhandedly	0,000,007	be_found	0,284,159
account_for	0,085,513	and_roadrunner	0,000,003	be_four	0,002,943
account_that	0,011,669	and_shaders	0,000,015	be_happy	0,054,320
achieve_a	0,014,392	and_so	0,881,129	be_held	0,043,088
achieved_an	0,000,851	and_tablets	0,000,652	be_merged	0,002,633
actual_power	0,000,681	and_the	9,342,102	be_more	0,203,851
admin_created	0,000,003	and_their	0,516,037	be_raw	0,000,193
advanced_institute	0,000,043	and_thin	0,012,580	be_right	0,032,920
advertising_new	0,000,022	and_thus	0,229,133	be_the	0,809,700
advised_to	0,010,119	and_two	0,142,601	be_well	0,061,438
affected_though	0,000,024	and_ultra	0,000,963	because_a	0,047,416
after_the	0,833,440	and_ulv	0,000,011	because_it	0,369,684
after_years	0,015,567	and_yet	0,381,477	because_of	0,506,531
ago_and	0,008,520	angry_like	0,000,070	before_they	0,106,230
ago_systems	0,000,001	any_bit	0,000,808	began_publishing	0,000,614
agreement_for	0,002,728	any_case	0,046,448	beginning_of	0,204,080
ahead_of	0,069,171	appears_that	0,030,918	behind_it	0,031,835
aics_in	0,000,008	apple_has	0,000,920	behind_the	0,199,162
alamos_national	0,000,432	architecture_for	0,005,470	being_consigned	0,000,184
all_amongst	0,000,148	are_all	0,208,340	benchmark_application	0,000,036
all_communication	0,002,881	are_allowed	0,024,670	benchmark_using	0,000,011

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

berkeley_national	0,000,318	combined_performance	0,000,027	does_not	1,195,076
best_of	0,099,084	combined_processors	0,000,001	doesn_t	0,397,801
best_silicon	0,000,002	comes_with	0,027,424	dominates_the	0,002,897
between_and	0,094,392	commented_on	0,004,934	don_t	2,521,353
birds_of	0,017,105	communication_concerning	0,000,056	down_by	0,061,853
bit_drivers	0,000,085	company_from	0,002,451	down_from	0,084,735
bits_are	0,006,452	company_sells	0,000,359	down_the	0,485,470
blame_for	0,009,413	company_would	0,003,373	dozen_more	0,001,782
blog_about	0,000,098	compared_to	0,098,750	dreaming_of	0,007,889
boss_even	0,000,007	comparison_to	0,010,836	drivers_for	0,005,074
boss_started	0,000,013	compatible_with	0,030,109	dual_core	0,001,278
branch_of	0,058,652	competitor_out	0,000,032	due_to	0,394,650
branch_under	0,000,120	competitors_like	0,000,054	eagerly_awaiting	0,000,386
brand_new	0,008,780	competitors_with	0,000,226	easily_leaves	0,000,012
break_the	0,042,297	compiled_by	0,024,976	edition_of	0,054,351
bridge_chip	0,000,248	completely_cut	0,000,697	efficiency_is	0,002,767
bridge_generation	0,000,002	computational_science	0,000,195	efficient_systems	0,000,103
bridge_in	0,003,081	computer_also	0,000,245	efficient_than	0,004,637
bridge_s	0,001,065	computer_fujitsu	0,000,003	empty_warehouses	0,000,027
bridges_and	0,003,126	computer_had	0,000,588	end_of	0,749,428
build_out	0,000,447	computer_is	0,015,382	energy_efficiency	0,002,491
built_k	0,000,011	computer_it	0,000,375	energy_efficient	0,002,080
business_sector	0,000,499	computer_maintained	0,000,011	energy_research	0,000,389
but_ibm	0,000,071	computing_center	0,000,402	english_town	0,000,416
but_it	0,960,416	concerning_the	0,119,039	enough_of	0,047,568
but_only	0,146,519	conference_in	0,005,187	entry_level	0,002,813
but_the	1,651,981	confirmed_to	0,001,727	epiphany_e	0,000,003
but_things	0,004,887	confuse_the	0,004,852	equipping_its	0,000,024
but_was	0,088,958	confusion_with	0,002,152	erich_strohmaier	0,000,001
but_why	0,055,108	consigned_to	0,005,076	es_interface	0,000,012
by_admin	0,000,030	consistent_fashion	0,000,134	ethernet_is	0,001,408
by_hans	0,001,815	consuming_only	0,000,067	ethernet_ones	0,000,005
by_infiniband	0,000,003	consumption_of	0,016,468	even_if	0,355,933
by_marcel	0,000,780	continues_to	0,045,646	even_intel	0,000,023
by_now	0,032,046	contrary_to	0,079,099	even_more	0,121,667
by_the	5,410,280	copyright_c	0,018,356	ever_to	0,029,061
by_their	0,228,795	copyrights_on	0,000,054	every_list	0,000,177
by_this	0,270,380	core_comes	0,000,134	everything_is	0,059,594
by_total	0,001,433	core_processors	0,000,323	expense_of	0,040,983
c_top	0,000,107	cores_and	0,000,452	extra_time	0,001,932
calculations_per	0,000,083	cores_compatible	0,000,001	failed_attempt	0,000,512
called_jaguar	0,000,001	costs_and	0,025,083	family_with	0,005,876
can_be	2,333,780	could_feature	0,000,022	far_as	0,351,348
can_currently	0,000,366	could_solve	0,001,174	fast_as	0,042,291
can_rejoice	0,000,757	cray_xt	0,000,001	fatal_mistake	0,001,052
captured_the	0,007,220	current_version	0,006,319	favor_of	0,058,836
card_geforce	0,000,003	currently_be	0,000,450	feather_session	0,000,006
cards_at	0,000,898	cut_us	0,001,156	feature_a	0,001,518
caused_by	0,183,577	damper_the	0,000,007	few_more	0,023,470
causes_confusion	0,000,225	data_behind	0,000,156	first_available	0,001,132
cell_processors	0,000,007	debut_in	0,002,912	first_ever	0,000,913
center_no	0,000,103	delay_should	0,000,244	first_supercomputer	0,000,015
changed_its	0,005,274	department_for	0,002,580	first_time	0,216,314
changes_to	0,064,433	derived_from	0,132,338	flood_in	0,001,157
charismatic_boss	0,000,001	desktop_graphics	0,000,030	followed_by	0,223,524
china_keeps	0,000,008	dethroned_when	0,000,001	for_almost	0,014,859
chinese_systems	0,000,015	diagram_with	0,000,855	for_bits	0,000,526
chip_s	0,000,700	dice_on	0,000,191	for_computational	0,000,966
chip_see	0,000,045	did_not	1,670,328	for_example	0,962,019
chips_with	0,000,357	didn_t	0,751,027	for_graphics	0,001,196
claim_the	0,011,886	disclosure_agreement	0,000,098	for_instance	0,235,623
clearly_the	0,012,868	discussion_at	0,001,244	for_mobile	0,006,556
clients_wait	0,000,025	disk_shortage	0,000,005	for_more	0,246,552
clock_speed	0,001,635	does_highlight	0,000,054	for_now	0,040,269

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

for_smartphone	0,000,069	have_any	0,122,283	interconnect_technology	0,000,052
for_some	0,318,233	have_instead	0,000,440	interest_include	0,000,452
for_the	6,068,591	have_its	0,020,579	interface_the	0,004,506
for_this	0,607,670	have_made	0,174,153	internal_diagram	0,000,003
for_workstation	0,000,161	have_to	1,015,441	internal_system	0,000,539
for_z	0,018,311	hd_graphics	0,000,005	into_account	0,043,027
former_amd	0,000,001	hd_instead	0,000,002	into_the	2,507,170
found_in	0,359,952	head_of	0,217,400	introduced_this	0,001,030
four_cpu	0,000,055	heavily_advertising	0,000,003	is_already	0,065,321
four_times	0,028,430	held_at	0,016,387	is_also	0,403,845
fourth_quarter	0,000,897	highest_total	0,000,166	is_being	0,106,203
france_round	0,000,009	highlight_a	0,001,250	is_compiled	0,004,841
from_all	0,190,278	his_boss	0,002,355	is_concerned	0,046,611
from_ibm	0,002,788	his_business	0,031,371	is_growing	0,011,848
from_its	0,150,980	his_printing	0,000,360	is_heavily	0,002,689
from_june	0,001,699	history_after	0,000,426	is_mainly	0,011,225
from_sending	0,001,755	hp_had	0,000,026	is_meanwhile	0,000,107
from_six	0,004,690	hp_is	0,000,410	is_mflops	0,000,003
from_systems	0,000,473	hp_starts	0,000,002	is_now	0,233,718
from_the	5,752,280	hp_with	0,000,049	is_one	0,380,342
full_build	0,000,056	hsun_huang	0,000,008	is_slightly	0,009,852
future_workstation	0,000,001	huang_s	0,000,219	is_sprouting	0,000,072
gave_us	0,022,499	ibm_kept	0,000,006	is_still	0,203,259
geforce_gtx	0,000,431	ibm_power	0,000,120	is_supposed	0,038,327
generation_before	0,000,650	ibm_stays	0,000,003	is_the	3,787,562
geometry_shaders	0,000,004	ibm_system	0,000,450	is_to	1,205,446
german_original	0,000,220	ibm_with	0,000,210	is_working	0,016,552
germany_nov	0,000,009	if_correct	0,000,438	it_achieved	0,000,626
gigabit_ethernet	0,004,371	if_mainly	0,000,019	it_and	0,247,935
gk_of	0,000,015	imagination_technologies	0,000,003	it_appears	0,098,620
going_to	0,883,715	implementations_or	0,000,109	it_could	0,159,276
gpu_designs	0,000,001	in_any	0,403,733	it_does	0,249,107
gpu_is	0,000,030	in_china	0,033,812	it_four	0,001,451
gpu_should	0,000,002	in_comparison	0,037,312	it_into	0,128,527
gpus_as	0,000,002	in_fact	0,432,361	it_market	0,000,068
gpus_to	0,000,012	in_favor	0,050,254	it_might	0,223,337
graphics_card	0,002,203	in_june	0,024,918	it_mildly	0,000,937
graphics_cards	0,001,035	in_kobe	0,000,134	it_must	0,283,805
graphics_core	0,000,028	in_market	0,002,413	it_reflects	0,003,629
graphics_of	0,000,239	in_nm	0,000,144	it_saves	0,003,282
graphics_performance	0,000,148	in_santa	0,002,699	it_takes	0,081,701
graphics_processors	0,000,131	in_seattle	0,002,188	it_tried	0,002,183
graphics_unit	0,000,020	in_second	0,005,520	it_will	0,666,736
graphics_will	0,000,161	in_shops	0,000,724	it_would	0,795,077
group_can	0,004,878	in_systems	0,002,869	its_competitors	0,001,464
growing_quickly	0,000,159	in_thailand	0,003,199	its_debut	0,000,310
grown_to	0,011,722	in_that	0,765,088	its_hat	0,000,203
gtx_and	0,000,031	in_the	9,999,999	its_large	0,003,078
had_already	0,138,276	increasing_its	0,001,684	its_lead	0,000,387
had_systems	0,000,026	industrial_customers	0,000,071	its_mind	0,002,484
had_taken	0,189,305	infiniband_based	0,000,002	its_nearest	0,000,918
half_a	0,150,952	infiniband_technology	0,000,015	its_number	0,001,688
happy_about	0,003,431	installed_at	0,002,834	its_own	0,306,786
hard_disk	0,034,802	instead_of	0,472,032	its_performance	0,003,895
has_changed	0,026,702	institute_for	0,013,011	its_position	0,010,966
has_cores	0,000,007	intel_already	0,000,006	its_widespread	0,000,723
has_grown	0,018,627	intel_clients	0,000,005	ivy_bridge	0,000,004
has_lost	0,023,102	intel_continues	0,000,008	ivy_overgrown	0,000,007
has_now	0,025,937	intel_doesn	0,000,013	jack_dongarra	0,000,026
has_produced	0,011,887	intel_has	0,000,440	jaguar_and	0,000,196
has_stabilized	0,000,431	intel_is	0,000,336	japan_and	0,006,311
hasn_t	0,057,135	intel_probably	0,000,002	japan_s	0,006,463
have_a	1,063,044	intel_processors	0,000,861	jen_hsun	0,000,006
have_annoyed	0,000,432	intel_s	0,002,736	jon_peddie	0,000,005

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

joshi_is	0,000,021	mantra_is	0,001,608	not_so	0,277,733
just_introduce	0,000,058	manufacturer_has	0,000,535	not_the	0,689,927
just_introduced	0,000,271	manufacturers_and	0,003,256	not_utilize	0,000,443
just_six	0,001,104	many_other	0,102,482	note_of	0,048,047
just_to	0,086,461	mark_on	0,008,574	notebook_manufacturers	0,000,057
k_computer	0,000,065	market_as	0,001,740	nothing_to	0,236,076
k_has	0,001,011	market_share	0,006,402	notice_and	0,004,167
keeps_increasing	0,000,105	matter_so	0,003,407	now_clearly	0,001,101
kept_its	0,001,990	maybe_with	0,000,556	now_followed	0,001,358
knightly_power	0,000,007	meager_annual	0,000,001	now_have	0,026,681
laboratory_no	0,000,027	mean_a	0,015,379	now_systems	0,000,005
langley_near	0,000,002	megawatt_mw	0,000,008	now_tracks	0,000,023
large_brother	0,000,005	merged_with	0,004,575	now_using	0,001,079
large_notebook	0,000,019	message_simply	0,000,119	now_wants	0,000,409
large_systems	0,001,138	mhz_of	0,000,266	now_with	0,021,869
largest_share	0,000,687	might_also	0,019,551	number_of	1,218,063
largest_u	0,000,231	might_be	0,651,259	numbers_might	0,000,338
larrabee_and	0,000,011	mind_on	0,011,535	nvdiia_chips	0,000,003
last_longer	0,002,198	mistake_to	0,006,886	nvdiia_gpus	0,000,013
last_minute	0,007,594	mobile_graphics	0,000,017	nvdiia_only	0,000,003
last_system	0,000,245	models_and	0,011,548	nvdiia_s	0,000,551
latest_list	0,000,121	monster_processor	0,000,001	nvdiia_tried	0,000,001
launch_date	0,000,424	months_ago	0,016,777	oak_ridge	0,000,915
lawrence_berkeley	0,000,494	more_cores	0,000,021	obviously_proud	0,000,040
lead_in	0,009,982	more_efficient	0,023,872	occupy_the	0,014,211
lead_of	0,004,242	more_gpu	0,000,001	of_all	1,340,017
lead_the	0,024,942	more_information	0,137,000	of_apple	0,003,224
leave_out	0,005,776	more_or	0,111,241	of_big	0,008,821
leaves_its	0,001,199	more_than	1,222,177	of_both	0,152,379
less_reasonably	0,000,018	more_that	0,014,393	of_clock	0,001,229
level_of	0,203,685	more_weeks	0,001,054	of_cray	0,000,218
level_to	0,011,031	most_efficient	0,008,811	of_equipping	0,000,211
like_a	0,981,190	most_energy	0,000,309	of_hd	0,000,688
linpack_benchmark	0,000,042	most_likely	0,059,196	of_hp	0,001,059
list_are	0,003,396	most_powerful	0,027,297	of_hpc	0,000,214
list_back	0,000,160	most_used	0,002,074	of_ibm	0,006,474
list_by	0,003,819	moved_up	0,004,584	of_interest	0,085,796
list_does	0,000,662	much_discussed	0,001,457	of_its	0,713,902
list_is	0,024,378	much_performance	0,000,097	of_ivybridge	0,000,002
list_moved	0,000,001	much_with	0,009,675	of_japan	0,013,106
list_of	0,298,622	must_have	0,481,946	of_kings	0,029,252
list_remain	0,000,008	mw_of	0,000,255	of_mannheim	0,000,218
list_the	0,018,351	mw_up	0,000,007	of_march	0,030,511
list_was	0,002,218	nasa_s	0,001,999	of_model	0,004,232
listed_at	0,003,381	national_energy	0,000,452	of_nersc	0,000,003
lists_press	0,000,001	national_laboratory	0,002,125	of_only	0,028,449
london_easily	0,000,003	near_london	0,001,374	of_other	0,203,352
longer_to	0,016,521	nearest_competitor	0,000,082	of_percent	0,009,264
los_alamos	0,001,614	need_some	0,009,756	of_performance	0,015,334
losers_are	0,000,138	nerves_seem	0,000,027	of_power	0,083,824
lost_track	0,001,891	new_barebones	0,000,001	of_service	0,060,874
low_and	0,029,568	new_models	0,001,066	of_supercomputers	0,000,119
low_voltage	0,003,427	new_nm	0,000,016	of_systems	0,011,660
lv_and	0,000,434	newest_edition	0,000,047	of_tennessee	0,003,987
machine_at	0,001,679	newest_list	0,000,004	of_the	9,999,999
made_changes	0,001,103	next_top	0,000,525	of_them	1,190,586
made_it	0,119,913	nine_systems	0,000,027	of_these	1,174,404
mainly_at	0,001,018	nm_atoms	0,000,002	of_this	2,014,150
mainly_derived	0,000,241	nm_production	0,000,001	of_todd	0,000,313
maintained_its	0,001,981	nm_structures	0,000,005	of_top	0,004,333
make_the	0,303,349	no_other	0,215,665	off_from	0,054,668
maker_of	0,011,327	no_turnover	0,000,012	off_their	0,023,424
makes_it	0,080,384	non_disclosure	0,000,314	old_opengl	0,000,001
makes_the	0,082,745	not_change	0,037,761	old_sandy	0,000,078

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

on_nm	0,000,022	performance_running	0,000,021	release_links	0,000,011
on_short	0,003,446	performance_slightly	0,000,056	remain_unchanged	0,002,813
on_the	8,210,281	pesky_problem	0,000,022	remaining_space	0,000,699
on_this	0,413,438	philosophized_about	0,000,103	reports_the	0,006,363
on_top	0,063,169	place_is	0,031,075	research_center	0,006,333
one_of	2,356,901	plants_need	0,000,245	research_scientific	0,000,029
one_or	0,246,978	plus_of	0,000,355	respective_owners	0,003,519
one_year	0,053,687	points_of	0,064,639	restructuring_the	0,000,434
ones_and	0,006,886	position_atop	0,000,032	resulted_in	0,055,002
ones_for	0,003,138	position_in	0,051,593	ridge_national	0,000,346
only_just	0,020,332	position_with	0,008,245	round_out	0,001,361
only_shader	0,000,001	possible_future	0,001,781	rumors_about	0,001,006
only_the	0,366,957	power_consumption	0,004,155	running_the	0,040,662
only_watts	0,000,107	power_efficiency	0,000,184	s_ames	0,000,044
opengl_es	0,000,183	power_from	0,009,118	s_barrier	0,000,145
opteron_family	0,000,001	power_processors	0,000,129	s_charismatic	0,000,081
optional_geometry	0,000,001	powerful_as	0,004,468	s_department	0,002,587
or_competitors	0,000,146	powerful_graphics	0,000,086	s_going	0,066,263
or_less	0,143,422	powerful_supercomputers	0,000,030	s_imagination	0,005,800
or_more	0,283,800	presented_new	0,000,124	s_launch	0,000,584
or_other	0,280,131	press_release	0,002,582	s_mark	0,001,866
or_quadrillion	0,000,013	previous_top	0,000,088	s_most	0,037,833
orders_in	0,005,929	previously_published	0,001,325	s_not	0,214,674
org_about	0,000,244	printing_group	0,000,024	s_one	0,041,230
org_all	0,000,068	probably_took	0,000,785	s_performance	0,008,698
org_blog	0,000,020	process_with	0,006,977	s_plants	0,000,222
org_contributors	0,000,008	processing_cores	0,000,011	s_processor	0,000,609
org_copyright	0,000,118	processor_power	0,000,383	s_six	0,002,443
org_http	0,001,236	processors_for	0,000,710	s_still	0,020,786
org_japan	0,000,001	processors_has	0,000,115	s_super	0,000,557
org_lists	0,000,087	processors_of	0,000,452	s_than	0,001,400
org_tags	0,000,002	processors_or	0,000,475	said_top	0,000,006
original_in	0,003,299	processors_will	0,000,374	sales_went	0,000,050
originate_from	0,005,350	processors_with	0,000,808	santa_clara	0,002,522
other_accelerators	0,000,016	produce_enough	0,000,715	saves_space	0,000,249
other_client	0,001,061	produced_by	0,111,305	sc_supercomputing	0,000,001
other_company	0,002,792	produced_in	0,028,595	scheduled_dates	0,000,059
other_developments	0,000,477	production_process	0,002,615	scheduled_for	0,004,399
other_points	0,006,830	proud_presented	0,000,002	scientific_computing	0,000,385
other_recent	0,000,800	provide_the	0,060,183	second_place	0,012,519
other_things	0,149,225	publicly_and	0,001,595	second_restructuring	0,000,004
other_top	0,000,515	published_message	0,000,037	sector_is	0,002,100
out_of	2,102,722	published_on	0,004,000	see_picture	0,000,123
out_that	0,097,419	published_six	0,000,180	seem_a	0,011,192
out_the	0,392,527	publishing_the	0,002,848	seem_to	0,304,839
out_what	0,049,178	pull_an	0,000,355	sell_off	0,000,815
over_a	0,186,783	put_it	0,137,860	selling_the	0,004,674
own_workstations	0,000,062	q_with	0,000,945	sells_more	0,000,127
owned_by	0,024,661	quadrillion_calculations	0,000,001	sending_us	0,001,005
pack_four	0,000,029	quarter_of	0,071,336	service_so	0,000,875
page_are	0,002,658	quickly_even	0,000,106	session_to	0,004,230
partially_built	0,000,075	quickly_pull	0,000,047	sgi_altix	0,000,006
partner_is	0,002,312	quiet_smile	0,001,369	shader_model	0,000,044
per_second	0,027,498	rankings_did	0,000,001	shader_units	0,000,046
percent_and	0,005,164	raw_in	0,000,383	shaders_and	0,000,085
percent_are	0,002,008	reasonable_prices	0,000,527	shaders_of	0,000,005
percent_behind	0,000,032	receive_these	0,001,629	share_by	0,000,638
percent_compared	0,000,306	received_because	0,000,118	share_is	0,002,305
percent_of	0,120,222	recent_very	0,000,047	share_of	0,053,548
percent_six	0,000,005	reduces_costs	0,000,182	shop_counters	0,000,012
percent_the	0,000,710	reflects_will	0,000,026	short_notice	0,003,035
performance_just	0,000,086	refrained_from	0,007,982	shortage_caused	0,000,034
performance_level	0,000,757	regarding_some	0,000,306	should_be	1,672,542
performance_of	0,071,689	rejoice_over	0,002,550	should_make	0,036,579

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

showed_no	0,015,435	system_interconnect	0,000,114	the_failed	0,004,910
silent_movie	0,000,178	system_is	0,088,152	the_first	2,329,965
simon_of	0,000,930	system_on	0,011,363	the_flood	0,029,767
simply_didn	0,000,797	system_that	0,043,131	the_former	0,221,405
since_we	0,048,018	system_took	0,000,182	the_fourth	0,146,629
sites_http	0,000,226	system_with	0,019,269	the_german	0,112,121
six_months	0,057,033	systems_account	0,000,031	the_gk	0,000,204
six_or	0,018,842	systems_and	0,037,416	the_gpu	0,000,373
slightly_ahead	0,000,270	systems_at	0,002,444	the_hard	0,058,506
slightly_down	0,000,163	systems_has	0,002,426	the_head	0,351,882
slip_us	0,000,048	systems_in	0,019,594	the_highest	0,232,989
small_town	0,008,783	systems_include	0,001,524	the_hp	0,002,709
smaller_competitors	0,000,060	systems_on	0,005,766	the_it	0,015,678
smile_about	0,000,453	systems_showed	0,000,050	the_ivy	0,003,947
smile_and	0,013,994	systems_to	0,014,380	the_japanese	0,034,201
so_much	0,596,315	systems_use	0,003,125	the_k	0,018,597
so_super	0,000,051	t_have	0,310,438	the_large	0,098,324
so_that	0,967,541	t_made	0,004,925	the_largest	0,096,728
soc_market	0,000,001	t_matter	0,035,378	the_last	0,907,982
socs_with	0,000,005	t_mean	0,058,191	the_latest	0,071,775
solve_a	0,005,243	t_need	0,068,927	the_lead	0,031,236
some_extra	0,003,615	t_suit	0,002,482	the_level	0,106,421
some_of	0,804,143	tablets_with	0,000,323	the_linpack	0,000,061
some_topics	0,000,498	taken_the	0,054,187	the_list	0,186,501
some_weeks	0,008,400	takes_longer	0,002,091	the_losers	0,001,838
something_to	0,133,724	taking_into	0,009,349	the_low	0,083,288
somewhere_between	0,005,052	talk_about	0,094,020	the_mantra	0,005,660
source_url	0,000,184	technology_systems	0,000,426	the_market	0,095,430
southern_english	0,000,094	technology_with	0,001,432	the_mobile	0,016,810
speaking_of	0,059,838	th_edition	0,012,002	the_monster	0,019,921
speed_while	0,000,216	th_spot	0,000,039	the_most	1,657,633
spot_in	0,012,411	than_all	0,046,144	the_much	0,013,288
spot_increases	0,000,009	than_gigabit	0,000,011	the_national	0,151,816
sprouting_and	0,000,156	than_half	0,035,853	the_new	0,926,541
stabilized_for	0,000,091	than_megawatt	0,000,004	the_newest	0,008,500
standard_for	0,018,223	thanks_to	0,065,227	the_next	0,786,438
standard_linpack	0,000,004	that_apple	0,000,928	the_no	0,019,421
start_selling	0,000,262	that_dominates	0,000,682	the_numbers	0,040,629
starts_a	0,004,782	that_his	0,375,810	the_oak	0,013,568
stay_atop	0,000,016	that_intel	0,000,359	the_old	0,952,763
stays_well	0,000,033	that_is	2,011,956	the_one	0,595,493
stem_from	0,004,292	that_knightly	0,000,071	the_ones	0,040,187
still_hasn	0,000,374	that_makes	0,066,115	the_opengl	0,001,708
still_in	0,087,925	that_the	4,586,627	the_overall	0,059,144
still_the	0,058,933	that_was	0,751,632	the_partially	0,001,885
still_unclear	0,001,129	the_and	0,001,585	the_pc	0,016,839
stocks_at	0,000,342	the_battery	0,017,272	the_petaflop	0,000,001
strohmaier_and	0,000,001	the_beginning	0,307,323	the_powervr	0,000,003
structures_by	0,001,034	the_best	0,644,060	the_previous	0,169,418
structures_that	0,008,405	the_brand	0,007,708	the_processors	0,001,959
suit_them	0,001,101	the_chinese	0,054,618	the_remaining	0,077,274
super_as	0,000,028	the_chip	0,006,524	the_riken	0,000,026
super_chip	0,000,020	the_company	0,217,720	the_sales	0,025,980
supercomputer_to	0,000,030	the_competition	0,011,160	the_sc	0,002,166
supercomputers_are	0,000,065	the_current	0,321,833	the_share	0,012,470
supercomputers_in	0,000,066	the_damper	0,000,536	the_shop	0,040,101
supercomputers_on	0,000,010	the_data	0,363,360	the_silent	0,031,650
supercomputing_conference	0,000,016	the_delays	0,001,967	the_small	0,177,471
supercomputing_sites	0,000,007	the_desktop	0,022,868	the_smaller	0,037,673
supposed_to	0,151,434	the_discussion	0,043,011	the_soc	0,000,552
surely_doesn	0,000,031	the_dual	0,008,946	the_southern	0,077,512
surely_say	0,000,349	the_end	0,781,246	the_spotlight	0,001,791
system_at	0,006,135	the_entry	0,035,534	the_standard	0,145,210
system_called	0,001,741	the_expense	0,045,348	the_system	0,384,591

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

the_systems	0,017,124	to_produce	0,134,790	use_processors	0,000,026
the_th	0,305,674	to_provide	0,213,261	used_internal	0,000,047
the_top	0,323,395	to_put	0,273,680	user_of	0,007,257
the_topic	0,021,963	to_quickly	0,011,062	using_intel	0,000,141
the_trends	0,001,605	to_reach	0,116,727	using_nvidia	0,000,018
the_twice	0,001,385	to_rumors	0,000,163	utilize_graphics	0,000,001
the_two	0,910,094	to_sell	0,065,219	vendors_are	0,001,742
the_unified	0,002,680	to_slip	0,014,245	very_large	0,050,376
the_university	0,129,650	to_smile	0,019,267	visit_www	0,003,592
the_upcoming	0,003,744	to_start	0,132,394	vyomesh_joshi	0,000,004
the_valley	0,102,185	to_stay	0,117,336	wafers_with	0,000,068
the_weaker	0,009,809	to_talk	0,159,059	wait_in	0,007,086
the_whims	0,001,574	to_the	9,999,999	wants_chips	0,000,001
the_world	1,612,409	to_their	0,591,025	was_at	0,252,174
their_clients	0,004,249	to_use	0,651,694	was_dethroned	0,000,415
their_orders	0,003,962	todd_bradley	0,000,001	was_listed	0,000,959
their_respective	0,038,221	took_note	0,001,664	was_published	0,030,342
their_sandy	0,000,086	took_the	0,208,546	was_the	1,675,634
them_have	0,019,917	top_editor	0,000,008	watt_one	0,000,003
there_are	1,254,177	top_just	0,000,149	watt_up	0,000,010
these_energy	0,000,304	top_list	0,000,315	we_are	0,966,046
these_use	0,000,459	top_now	0,000,113	we_began	0,012,868
they_are	1,716,107	top_rankings	0,000,040	weeks_to	0,010,783
they_have	0,673,615	top_spot	0,000,215	well_advised	0,001,906
they_will	0,343,345	top_supercomputers	0,000,013	well_ahead	0,001,083
thin_clients	0,000,915	top_supercomputing	0,000,003	went_down	0,069,838
things_are	0,112,774	top_system	0,000,056	what_a	0,258,257
things_because	0,002,414	top_systems	0,000,024	what_s	0,295,191
thirty_nine	0,005,894	top_u	0,000,128	what_you	0,408,207
this_architecture	0,002,691	topic_of	0,015,052	when_it	0,422,395
this_branch	0,005,005	total_combined	0,000,125	when_the	1,618,577
this_doesn	0,007,272	total_of	0,047,408	when_version	0,000,086
this_is	1,904,005	total_performance	0,000,108	which_had	0,408,413
this_page	0,043,716	total_power	0,000,885	which_have	0,136,072
this_summer	0,006,963	touchpad_and	0,000,014	which_resulted	0,005,834
this_surely	0,000,913	town_of	0,057,622	which_the	1,138,906
this_with	0,029,390	track_by	0,000,976	which_will	0,127,215
this_year	0,049,098	tracks_actual	0,000,002	while_everything	0,000,732
three_dimensional	0,021,626	trademarks_and	0,002,036	while_the	0,439,496
thus_it	0,038,600	transistors_that	0,000,142	whims_of	0,001,586
time_since	0,020,093	translation_of	0,046,149	who_will	0,108,983
time_this	0,011,683	trends_it	0,000,069	whole_received	0,000,013
times_as	0,028,682	tried_to	0,309,615	why_should	0,113,601
tiny_ones	0,000,129	twice_as	0,026,275	widespread_use	0,005,602
to_accelerate	0,004,777	twice_yearly	0,000,254	will_be	1,980,103
to_achieve	0,090,627	two_chinese	0,000,484	will_feature	0,000,328
to_an	0,456,465	two_use	0,000,135	will_have	0,334,772
to_and	0,133,104	ultra_low	0,000,448	will_lead	0,020,716
to_be	7,073,486	unchanged_from	0,001,048	will_most	0,008,136
to_blame	0,033,253	under_the	0,909,733	will_now	0,030,235
to_break	0,105,149	unified_shaders	0,000,001	will_only	0,041,551
to_claim	0,023,688	unit_with	0,001,740	will_pack	0,000,302
to_completely	0,005,995	university_of	0,218,659	windows_tablet	0,000,006
to_equip	0,002,526	unperturbed_by	0,000,310	with_a	3,432,860
to_history	0,005,065	up_from	0,098,411	with_cores	0,000,032
to_hp	0,000,656	up_to	0,838,320	with_empty	0,003,710
to_ibm	0,005,156	us_off	0,006,008	with_every	0,056,479
to_its	0,290,224	us_something	0,004,923	with_hd	0,000,262
to_jon	0,000,352	use_at	0,008,001	with_ivy	0,001,620
to_make	1,325,148	use_ati	0,000,008	with_kepler	0,000,083
to_manage	0,072,122	use_cell	0,000,192	with_more	0,086,599
to_many	0,065,058	use_gpus	0,000,001	with_nm	0,000,195
to_occupy	0,022,150	use_more	0,005,906	with_old	0,011,625
to_originate	0,002,282	use_nvidia	0,000,006	with_shader	0,000,016

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

with_six	0,009,227	won_t	0,451,325	xe_at	0,000,016
with_somewhere	0,000,093	working_on	0,048,331	xt_system	0,000,010
with_systems	0,001,003	workstation_and	0,001,940	year_ago	0,011,808
with_that	0,242,233	workstations_with	0,000,292	yearly_list	0,000,006
with_the	5,438,118	world_s	0,103,937	years_of	0,230,750
with_three	0,036,099	would_be	1,527,405	yet_due	0,000,207
with_touchpad	0,000,004	would_seem	0,063,058	you_don	0,442,948
with_us	0,166,448	would_surely	0,010,825		

The first half of next phrase-checked (with **Gamma** corpus) 3-gram wordlist contains unfamiliar 3-grams whereas the second half holds all familiar to **Gamma** 3-grams with their corresponding occurrences:

Overlapped lines, i.e. lines common for both files: **938**

Unfamiliar lines, i.e. lines from 1st file not encountered in 2nd file: **600**

a_certain_ivy	average_power_efficiency	cores_compatible_with
a_cray_xe	barebones_with_old	cpu_dice_on
a_giant_ceramic	based_systems_account	cray_xe_at
a_million_wafers	best_silicon_dice	cray_xt_system
a_strongly_downgraded	between_and_shaders	current_gpgpu_implementations
a_system_tool	big_intel_clients	current_ipad_powervr
a_tegra_all	bits_are_nothing	damper_the_it
a_top_system	blog_about_ivy	delivering_mflops_watt
a_windows_tablet	bluegene_q_with	department_for_workstation
a_wolf_speaking	boss_even_more	dethroned_when_the
about_a_tegra	boss_started_restructuring	dice_receive_these
about_ivy_bridges	brand_new_barebones	dimensional_finfet_transistors
about_the_tegra	brand_new_nm	dinosaur_vj_vyomesh
accelerators_up_from	break_the_petaflop	disclosure_agreement_for
account_that_intel	bridge_delay_should	discussed_cuda_operations
adapteva_would_surely	bridge_generation_before	discussion_at_sc
advanced_institute_for	bridge_processors_for	disk_shortage_caused
advertising_new_ultrabooks	bridge_s_launch	dominates_the_soc
advised_to_quickly	bridge_stocks_at	dongarra_of_the
ago_and_average	bridges_and_tegra	down_from_systems
ago_and_mflops	built_k_computer	downgraded_kepler_architecture
agreement_for_z	but_ibm_stays	dreaming_of_equipping
ahead_of_cray	by_admin_created	easily_leaves_its
aics_in_kobe	by_hans_meuer	editor_erich_strohmaier
all_amongst_other	by_infiniband_technology	efficiency_is_mflops
all_communication_concerning	by_marcel_sieslack	efficient_supercomputers_are
all_current_gpgpu	by_total_performance	energy_efficiency_classifications
all_systems_has	called_jaguar_and	energy_efficient_supercomputers
all_using_nvidia	captured_the_no	epiphany_e_k
altix_machine_at	certain_ivy_overgrown	equip_brand_new
amd_manufacturer_has	charismatic_boss_even	equip_tegra_soc
an_internal_diagram	china_keeps_increasing	equipping_its_own
an_ipad_competitor	chinese_systems_at	erich_strohmaier_and
an_sgi_altix	chinese_tianhe_a	es_interface_the
and_also_refrained	chip_s_launch	everything_is_sprouting
and_france_round	chip_see_picture	family_with_systems
and_horst_simon	chip_tegra_can	feather_session_to
and_knights_corner	chips_with_hd	feature_a_strongly
and_merrily_philosophized	combined_processors_will	finfet_transistors_that
and_mflops_watt	comes_with_somewhere	first_available_pictures
and_more_cores	compatible_with_shader	first_supercomputer_to
and_nvidia_launches	competitors_like_larrabee	flood_in_thailand
and_offhandedly_commented	competitors_qualcomm_percent	followed_by_infiniband
and_tegra_hitches	competitors_with_empty	for_bits_are
and_ulv_at	completely_cut_us	for_smartphone_chips
andreas_stiller_http	computational_science_aics	for_z_mainboards
annoyed_nvidia_s	computer_also_reports	former_amd_manufacturer
annual_plus_of	computer_fujitsu_captured	found_in_superphones
apple_has_changed	computer_it_achieved	four_cpu_dice
apple_publicly_and	computer_tops_petaflop	france_round_out
architecture_for_graphics	computing_center_no	from_mflops_watt
are_bluegene_q	consumption_of_supercomputers	fujitsu_captured_the
arrived_in_shops	contributors_stiller_published	future_workstation_tablets
as_accelerators_up	contributors_stiller_translation	geforce_core_comes
at_industrial_customers	copyright_c_top	geometry_shaders_and
atop_top_list	copyrights_on_this	giant_ceramic_module
attempt_and_nvidia	core_codename_valleyview	gigabit_ethernet_ones
attempt_with_touchpad	core_processors_of	gk_chip_see

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamma r.15 corpus (powered by: 9,181,275 1-grams; 124,669,942 2-grams; 477,829,381 3-grams; 879,557,846 4-grams);

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

globalfoundries_can_rejoice
gpgpu_implementations_or
gpu_is_supposed
gpu_should_be
gpus_as_accelerators
gpus_to_accelerate
graphics_core_codename
graphics_unit_with
gtx_and_also
hans_meuer_of
hard_disk_shortage
has_cores_admittedly
has_now_systems
have_annoyed_nvidia
head_of_hp
heavily_advertising_new
highest_total_power
his_boss_started
his_business_sector
his_printing_group
history_after_years
hitches_by_admin
horst_simon_of
hp_dinosaur_vj
hp_had_systems
hp_is_meanwhile
hp_maybe_with
hp_s_department
hp_starts_a
hp_with_systems
hsun_huang_s
huang_s_super
ibm_power_processors
ibm_stays_well
ibm_with_systems
implementations_or_competitors
in_nm_structures
in_superphones_and
infiniband_based_systems
infiniband_technology_with
intel_already_gave
intel_causes_confusion
intel_has_lost
intel_is_heavily
intel_probably_took
intel_s_plants
interconnect_technology_systems
interludes_behind_the
internal_diagram_with
internal_system_interconnect
introduce_optional_geometry
introduced_this_architecture
ipad_competitor_out
ipad_powervr_core
is_heavily_advertising
is_meanwhile_dreaming
is_mflops_watt
it_could_feature
it_market_as
its_competitors_qualcomm
its_large_brother
ivy_bridge_chip
ivy_bridge_delay
ivy_bridge_generation

ivy_bridge_processors
ivy_bridge_s
ivy_bridges_and
ivy_overgrown_bridge
ivybridge_still_hasn
jack_dongarra_of
jaguar_and_installed
japan_s_k
jim_zarafana_obviously
jon_peddie_research
joshi_is_being
just_introduce_optional
k_computer_also
k_computer_fujitsu
k_computer_had
k_computer_is
k_computer_it
k_computer_maintained
k_computer_tops
k_has_cores
kepler_graphics_card
kings_langley_near
langley_near_london
large_notebook_manufacturers
larrabee_and_knights
latest_list_the
launches_rumors_about
lead_in_systems
lead_of_todd
leaves_its_competitors
less_reasonably_scheduled
level_of_petaflop
like_larrabee_and
linpack_benchmark_application
linpack_benchmark_using
list_are_confirmed
list_by_admin
list_does_highlight
list_moved_up
list_remain_unchanged
list_was_listed
listed_at_position
lists_press_release
london_easily_leaves
los_alamos_no
low_voltage_variants
make_the_ultrabook
maker_of_powervr
marcel_sieslack_intel
meager_annual_plus
meanwhile_dreaming_of
merrily_philosophized_about
message_simply_didn
meuer_of_the
mflops_watt_one
mflops_watt_six
mflops_watt_unchanged
mflops_watt_up
million_wafers_with
mobile_graphics_will
mobile_ulp_geforce
models_and_merrily
monster_processor_power
months_ago_systems

more_gpu_designs
most_used_internal
movie_interludes_behind
much_discussed_cuda
mw_six_months
mw_up_from
national_laboratory_no
near_london_easily
nersc_lawrence_berkeley
new_barebones_with
new_kepler_graphics
new_nm_production
new_ultrabooks_with
newcomer_adapteva_would
newest_list_does
newest_list_was
next_top_list
nine_systems_use
nm_atoms_with
nm_production_process
nm_structures_by
nm_structures_that
not_utilize_graphics
notebook_manufacturers_happy
now_tracks_actual
now_using_intel
now_wants_chips
now_with_systems
nvidia_equip_tegra
nvidia_gpus_to
nvidia_launches_rumors
nvidia_only_just
nvidia_s_charismatic
obviously_proud_presented
of_big_intel
of_equipping_its
of_hd_graphics
of_ivybridge_still
of_kings_langley
of_nersc_lawrence
of_only_shader
of_petaflop_s
of_supercomputers_top
of_the_tegra
of_todd_bradley
off_their_sandy
offhandedly_commented_on
old_sandy_bridge
on_nm_atoms
on_top_supercomputing
ones_for_bits
only_just_introduce
only_shader_units
opengl_es_interface
opteron_family_with
optional_geometry_shaders
or_quadrillion_calculations
org_about_ivy
org_all_trademarks
org_blog_about
org_contributors_stiller
org_japan_s
org_lists_press
org_tags_whispers

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

other_company_sells
other_recent_very
overall_ivy_bridge
overgrown_bridge_in
pack_four_cpu
page_are_owned
partially_built_k
partner_is_working
pc_branch_under
percent_and_nvidia
performance_running_the
performance_slightly_ahead
petaflop_s_barrier
petaflop_s_on
petaflop_s_one
petaflop_s_performance
petaflop_s_six
petaflop_s_than
petaflop_s_to
philosophized_about_possible
possible_future_workstation
power_from_ibm
power_processors_has
powerful_graphics_core
powervr_graphics_of
presented_new_models
press_release_links
previous_top_just
previously_published_message
printing_group_can
probably_took_note
processor_epiphany_e
processor_power_from
processors_has_stabilized
processors_with_six
produced_in_nm
proud_presented_new
provide_the_processors
publicly_and_offhandedly
published_message_simply
pull_an_ipad
quadrillion_calculations_per
raw_in_santa
reasonably_scheduled_dates
receive_these_energy
regarding_some_topics
restructuring_attempt_and
riken_advanced_institute
s_charismatic_boss
s_imagination_technologies
s_k_computer
s_launch_date
s_performance_running
s_plants_need
s_processor_epiphany
s_super_chip
s_than_gigabit
said_top_editor
sandy_bridge_processors
sandy_bridge_stocks
science_aics_in
second_restructuring_attempt
selling_the_dual
sells_more_gpu

sgi_altix_machine
shaders_and_is
shaders_of_model
share_by_total
share_is_concerned
should_nvidia equip
showed_no_turnover
sieslack_intel_causes
silent_movie_interludes
silicon_dice_receive
simon_of_nersc
so_super_as
socs_with_cores
southern_english_town
sparc_processing_cores
speaking_of_powervr
sprouting_and_leafing
stabilized_for_now
started_restructuring_the
stay_atop_top
stays_well_ahead
still_in_second
stiller_published_on
stiller_translation_of
stocks_at_reasonable
strohmaier_and_horst
strongly_downgraded_kepler
structures_by_globalfoundries
super_as_far
super_chip_tegra
supercomputer_to_achieve
supercomputers_are_bluegene
supercomputers_in_a
supercomputers_top_now
supercomputing_sites_http
superphones_and_tablets
system_called_jaguar
system_interconnect_technology
systems_include_cielo
systems_use_gpus
t_by_marcel
tablets_with_us
tags_whispers_copyright
taken_the_no
technology_with_systems
tegra_all_amongst
tegra_can_currently
tegra_combined_processors
tegra_gpu_should
tegra_hitches_by
tegra_hitches_links
tegra_in_comparison
tegra_socs_with
teraflop_s_mark
teraflop_s_six
teraflops_at_mhz
th_spot_increases
than_all_current
than_megawatt_mw
that_apple_publicly
that_knightly_power
the_chinese_tianhe
the_current_atoms
the_current_ipad

the_former_amd
the_gk_chip
the_hp_dinosaur
the_ivy_bridge
the_japanese_tsubame
the_list_moved
the_mobile_ulp
the_monster_processor
the_new_kepler
the_overall_ivy
the_pc_branch
the_petaflop_s
the_powervr_graphics
the_riken_advanced
the_tegra_gpu
the_tegra_in
the_tianhe_a
the_ultrabook_manufacturers
the_unified_shaders
the_upcoming_tegra
the_valley_ultimately
the_weaker_graphics
their_clients_wait
their_sandy_bridge
these_energy_efficiency
these_use_nvidia
thirty_nine_systems
three_dimensional_finfet
tianhe_a_system
to_accelerate_computation
to_equip_brand
to_history_after
to_jon_peddie
top_editor_erich
top_just_six
top_now_tracks
top_rankings_did
top_supercomputers_on
top_supercomputing_sites
top_system_is
top_systems_showed
tops_petaflop_s
total_combined_performance
total_performance_slightly
touchpad_and_webos
town_of_ivybridge
tracks_actual_power
trends_it_reflects
twice_yearly_list
two_use_ati
two_use_cell
ulp_geforce_core
ultrabook_manufacturers_and
ultrabooks_with_ivy
ulp_at_the
unchanged_from_june
unified_shaders_of
unit_with_kepler
up_from_mflops
up_from_systems
upcoming_tegra_combined
use_at_industrial
use_ati_radeon
use_cell_processors

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

use_gpus_as
use_nvidia_chips
used_internal_system
user_of_hpc
using_nvidia_gpus
utilize_graphics_processors
variants_lv_and
vj_vyomesh_joshi
voltage_variants_lv
vyomesh_joshi_is
wait_in_favor
wants_chips_with
was_dethroned_when
watt_one_year
watt_six_months
watt_up_from
weaker_graphics_performance
which_the_tegra
whims_of_big
whispers_copyright_c
whole_received_because
why_should_nvidia
will_pack_four
with_cores_compatible
with_empty_warehouses
with_every_list
with_hd_instead
with_ivy_bridge
with_kepler_architecture
with_nm_structures
with_that_knightly
with_touchpad_and
wolf_speaking_of
working_on_nm
workstation_and_thin
workstation_tablets_with
workstations_with_that
xe_at_los
xt_system_called
yearly_list_remain
zarafana_obviously_proud
zarafana_won_t

a_birds_of	0,000,007	and_is_already	0,000,553	be_happy_about	0,000,664
a_consistent_fashion	0,000,089	and_is_mainly	0,000,251	be_held_at	0,001,876
a_cray_xt	0,000,001	and_is_now	0,010,044	be_merged_with	0,000,423
a_dozen_more	0,001,467	and_is_supposed	0,000,842	be_more_efficient	0,001,734
a_fatal_mistake	0,000,492	and_jack_dongarra	0,000,005	be_raw_in	0,000,007
a_feather_session	0,000,006	and_makes_the	0,005,211	be_the_topic	0,000,150
a_few_more	0,022,313	and_now_wants	0,000,064	be_well_advised	0,000,712
a_full_build	0,000,024	and_the_japanese	0,001,284	because_a_previously	0,000,004
a_graphics_unit	0,000,001	and_the_ones	0,001,357	because_it_saves	0,000,223
a_market_share	0,000,161	and_the_trends	0,000,052	because_of_apple	0,000,004
a_meager_annual	0,000,001	and_their_clients	0,000,451	because_of_the	0,185,824
a_non_disclosure	0,000,065	and_thin_clients	0,000,061	before_they_are	0,008,845
a_number_of	0,318,258	and_thus_it	0,008,809	began_publishing_the	0,000,081
a_partner_is	0,000,149	and_two_use	0,000,004	beginning_of_march	0,000,526
a_performance_level	0,000,059	and_ultra_low	0,000,026	behind_it_and	0,000,841
a_performance_of	0,001,502	angry_like_a	0,000,002	behind_the_shop	0,000,359
a_pesky_problem	0,000,006	any_bit_drivers	0,000,003	being_consigned_to	0,000,171
a_powerful_graphics	0,000,014	appears_that_apple	0,000,001	berkeley_national_laboratory	0,000,274
a_previously_published	0,000,041	are_all_using	0,000,091	best_of_the	0,010,889
a_quarter_of	0,043,606	are_allowed_to	0,013,641	birds_of_a	0,000,876
a_quiet_smile	0,000,924	are_confirmed_to	0,000,040	bit_drivers_for	0,000,011
a_second_restructuring	0,000,002	are_eagerly_awaiting	0,000,011	blame_for_the	0,001,974
a_system_with	0,002,917	are_not_so	0,018,880	branch_of_the	0,027,086
a_total_of	0,033,707	are_nothing_to	0,001,439	branch_under_the	0,000,044
a_user_of	0,001,044	are_now_followed	0,000,017	bridge_chip_s	0,000,002
a_whole_received	0,000,003	are_now_using	0,000,392	bridge_in_the	0,000,942
able_to_manage	0,002,799	are_owed_by	0,001,843	build_out_that	0,000,004
about_possible_future	0,000,032	are_rumors_about	0,000,018	business_sector_is	0,000,023
about_the_top	0,000,305	are_supposed_to	0,016,543	but_it_could	0,006,546
about_when_it	0,000,220	are_the_smaller	0,000,124	but_only_the	0,009,613
according_to_an	0,003,787	as_a_user	0,003,388	but_the_whims	0,000,015
according_to_jon	0,000,011	as_a_whole	0,050,751	but_things_are	0,000,790
according_to_rumors	0,000,012	as_far_as	0,196,225	but_was_dethroned	0,000,007
according_to_the	0,303,387	as_fast_as	0,038,290	but_why_should	0,007,719
account_for_almost	0,000,226	as_for_the	0,057,195	by_the_and	0,000,055
achieve_a_performance	0,000,035	as_his_business	0,000,180	by_the_flood	0,001,169
achieved_an_impressive	0,000,018	as_its_nearest	0,000,008	by_their_respective	0,001,406
actual_power_consumption	0,000,018	as_much_performance	0,000,024	calculations_per_second	0,000,059
after_the_failed	0,000,124	as_powerful_as	0,002,369	can_be_merged	0,000,460
after_years_of	0,006,555	as_the_current	0,003,502	can_currently_be	0,000,108
ahead_of_both	0,000,026	as_the_market	0,000,801	can_rejoice_over	0,000,016
ahead_of_japan	0,000,013	at_los_alamos	0,000,356	card_geforce_gtx	0,000,002
amos_national_laboratory	0,000,349	at_mhz_of	0,000,002	cards_at_the	0,000,264
all_trademarks_and	0,000,776	at_nasa_s	0,000,071	caused_by_the	0,036,056
allowed_to_talk	0,000,440	at_position_in	0,000,296	causes_confusion_with	0,000,011
almost_twice_as	0,000,579	at_reasonable_prices	0,000,328	changed_its_mind	0,000,167
already_arrived_in	0,000,096	at_the_beginning	0,082,906	changes_to_their	0,000,466
already_gave_us	0,000,004	at_the_end	0,246,531	claim_the_th	0,000,011
already_working_on	0,000,202	at_the_expense	0,026,755	clearly_the_no	0,000,002
also_one_of	0,005,168	at_the_last	0,038,212	clients_wait_in	0,000,003
also_refrained_from	0,000,030	at_the_national	0,004,862	clock_speed_while	0,000,005
also_reports_the	0,000,121	at_the_oak	0,000,294	combined_performance_of	0,000,007
also_stem_from	0,000,073	at_the_riken	0,000,002	commented_on_the	0,001,582
although_the_top	0,000,040	at_the_sc	0,000,010	communication_concerning_the	0,000,010
amd_opteron_family	0,000,001	atoms_with_a	0,000,088	company_from_the	0,000,527
ames_research_center	0,000,270	atop_the_newest	0,000,001	company_sells_more	0,000,006
amongst_other_things	0,001,771	average_power_consumption	0,000,036	company_would_be	0,000,698
an_ibm_system	0,000,059	awaiting_the_silent	0,000,005	compared_to_hp	0,000,002
and_average_power	0,000,032	back_in_that	0,000,809	compared_to_ibm	0,000,006
and_be_four	0,000,002	battery_last_longer	0,000,002	comparison_to_the	0,003,335
and_copyrights_on	0,000,008	be_able_to	0,406,236	competitor_out_of	0,000,027
and_for_the	0,093,439	be_affected_though	0,000,006	compiled_by_hans	0,000,002
and_has_now	0,000,906	be_found_in	0,101,233	computer_had_taken	0,000,002
and_installed_at	0,000,077	be_four_times	0,000,164	computer_is_also	0,000,079

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

computer_is_the	0,000,432	few_more_weeks	0,000,207	in_any_case	0,044,921
computer_maintained_its	0,000,002	first_ever_to	0,000,051	in_comparison_to	0,007,739
concerning_the_new	0,000,340	first_time_since	0,007,963	in_favor_of	0,047,355
conference_in_seattle	0,000,011	followed_by_the	0,044,735	in_market_share	0,000,192
confirmed_to_use	0,000,001	for_almost_twice	0,000,004	in_santa_clara	0,000,240
confuse_the_one	0,000,009	for_computational_science	0,000,019	in_second_place	0,000,212
confusion_with_the	0,000,518	for_graphics_cards	0,000,030	in_systems_and	0,000,107
consigned_to_history	0,000,019	for_mobile_graphics	0,000,001	in_that_the	0,011,936
consuming_only_watts	0,000,002	for_more_information	0,108,292	in_the_fourth	0,019,203
consumption_of_a	0,000,896	for_now_with	0,000,047	in_the_gk	0,000,018
continues_to_provide	0,000,359	for_some_weeks	0,003,245	in_the_latest	0,002,577
contrary_to_many	0,000,077	for_the_delays	0,000,061	in_the_previous	0,048,144
core_comes_with	0,000,118	for_the_ivy	0,000,034	in_the_southern	0,013,163
cores_and_more	0,000,001	for_the_largest	0,001,281	increasing_its_number	0,000,008
costs_and_makes	0,000,012	for_the_much	0,000,458	installed_at_the	0,000,993
could_feature_a	0,000,005	for_the_powervr	0,000,001	instead_of_hd	0,000,010
could_solve_a	0,000,040	for_this_summer	0,000,122	instead_of_only	0,001,023
currently_be_found	0,000,012	for_workstation_and	0,000,018	institute_for_computational	0,000,008
cut_us_off	0,000,598	four_times_as	0,002,530	intel_continues_to	0,000,007
data_behind_it	0,000,002	from_all_communication	0,000,354	intel_doesn_t	0,000,013
debut_in_the	0,000,552	from_its_large	0,000,098	interface_the_standard	0,000,016
delay_should_make	0,000,002	from_sending_us	0,000,017	into_account_that	0,001,278
derived_from_the	0,044,652	from_six_months	0,000,179	into_the_spotlight	0,000,124
desktop_graphics_card	0,000,002	from_the_damper	0,000,007	is_a_cray	0,000,002
diagram_with_more	0,000,002	from_the_old	0,019,259	is_already_working	0,000,108
dice_on_a	0,000,034	from_the_small	0,003,358	is_also_one	0,002,780
did_not_change	0,004,647	full_build_out	0,000,003	is_being_consigned	0,000,006
didn_t_suit	0,000,435	gave_us_something	0,000,084	is_compiled_by	0,000,160
does_highlight_a	0,000,010	geforce_gtx_and	0,000,017	is_growing_quickly	0,000,044
does_not_utilize	0,000,151	generation_before_they	0,000,007	is_mainly_derived	0,000,088
doesn_t_have	0,023,238	german_original_in	0,000,008	is_now_clearly	0,000,148
doesn_t_matter	0,022,970	gigabit_ethernet_is	0,000,125	is_one_of	0,165,292
doesn_t_mean	0,012,674	gk_of_the	0,000,004	is_slightly_down	0,000,004
don_t_need	0,046,468	going_to_occupy	0,000,076	is_sprouting_and	0,000,005
down_from_six	0,000,008	graphics_card_geforce	0,000,002	is_still_the	0,009,712
down_the_valley	0,003,895	graphics_cards_at	0,000,001	is_supposed_to	0,030,709
dozen_more_or	0,000,028	graphics_of_the	0,000,081	is_the_chinese	0,000,240
drivers_for_the	0,000,880	graphics_performance_of	0,000,005	is_the_first	0,057,261
dual_core_processors	0,000,164	graphics_processors_or	0,000,002	is_to_blame	0,002,152
due_to_its	0,007,243	graphics_will_only	0,000,002	is_working_on	0,001,920
eagerly_awaiting_the	0,000,119	group_can_be	0,001,540	it_achieved_an	0,000,007
edition_of_the	0,018,371	growing_quickly_even	0,000,001	it_and_the	0,015,949
efficient_systems_on	0,000,002	had_already_arrived	0,000,894	it_appears_that	0,027,495
efficient_than_all	0,000,003	had_taken_the	0,019,928	it_could_solve	0,000,022
end_of_march	0,001,662	half_a_dozen	0,039,039	it_does_not	0,152,436
energy_efficient_systems	0,000,007	has_changed_its	0,000,599	it_four_times	0,000,291
energy_research_scientific	0,000,008	has_grown_to	0,001,686	it_into_the	0,036,401
english_town_of	0,000,042	has_lost_track	0,000,132	it_might_also	0,001,956
enough_of_them	0,002,396	has_produced_by	0,000,032	it_must_have	0,041,892
entry_level_of	0,000,031	has_stabilized_for	0,000,011	it_reflects_will	0,000,001
ethernet_is_still	0,000,018	hasn_t_made	0,000,533	it_s_imagination	0,000,026
even_if_mainly	0,000,008	have_a_few	0,008,903	it_s_processor	0,000,003
even_intel_has	0,000,003	have_any_bit	0,000,006	it_s_still	0,007,617
even_more_that	0,000,103	have_instead_of	0,000,130	it_saves_space	0,000,033
ever_to_break	0,000,047	have_its_debut	0,000,002	it_takes_longer	0,000,601
expense_of_the	0,010,931	have_made_changes	0,000,309	it_tried_to	0,001,542
extra_time_this	0,000,001	have_to_equip	0,000,041	it_will_pack	0,000,006
failed_attempt_with	0,000,001	hd_instead_of	0,000,001	it_would_seem	0,037,577
far_as_the	0,057,092	highlight_a_number	0,000,035	its_debut_in	0,000,113
fast_as_the	0,003,749	ibm_kept_its	0,000,003	its_lead_in	0,000,032
fatal_mistake_to	0,000,069	ibm_system_that	0,000,002	its_mind_on	0,000,045
favor_of_apple	0,000,001	if_mainly_at	0,000,005	its_nearest_competitor	0,000,031
favor_of_the	0,017,540	in_a_consistent	0,001,959	its_number_of	0,000,376
feature_a_graphics	0,000,001	in_a_meager	0,000,018	its_own_workstations	0,000,001

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

its_position_atop	0,000,002	most_powerful_supercomputers	0,000,018	on_short_notice	0,000,759
its_widespread_use	0,000,299	moved_up_to	0,000,987	on_the_latest	0,001,385
japan_and_france	0,000,029	much_with_the	0,002,020	on_the_linpack	0,000,003
jen_hsun_huang	0,000,006	must_have_annoyed	0,000,024	on_the_list	0,009,363
just_introduced_this	0,000,004	mw_of_power	0,000,054	on_the_newest	0,000,152
just_six_months	0,000,210	nasa_s_ames	0,000,028	on_the_weaker	0,000,187
just_to_claim	0,000,029	national_energy_research	0,000,018	on_this_page	0,004,307
keeps_increasing_its	0,000,004	need_some_extra	0,000,095	one_of_the	1,241,115
kept_its_lead	0,000,003	nerves_seem_to	0,000,015	one_or_other	0,009,455
last_system_on	0,000,001	new_models_and	0,000,065	one_year_ago	0,000,481
lawrence_berkeley_national	0,000,296	newest_edition_of	0,000,028	ones_and_is	0,000,026
lead_the_discussion	0,000,029	no_other_company	0,000,288	only_just_introduced	0,000,006
leave_out_what	0,000,034	non_disclosure_agreement	0,000,091	only_the_best	0,001,562
level_of_performance	0,001,169	not_so_super	0,000,005	or_competitors_like	0,000,001
level_to_the	0,001,651	not_the_brand	0,000,038	or_less_reasonably	0,000,003
like_a_wolf	0,001,290	note_of_this	0,001,335	or_more_cores	0,000,015
list_back_in	0,000,009	nothing_to_be	0,019,310	or_other_accelerators	0,000,003
list_is_compiled	0,000,022	notice_and_now	0,000,001	or_other_client	0,000,082
list_of_the	0,027,291	now_clearly_the	0,000,020	orders_in_favor	0,000,004
list_the_th	0,000,004	now_followed_by	0,000,115	org_copyright_c	0,000,018
list_the_top	0,000,072	now_have_to	0,001,791	original_in_c	0,000,002
list_was_published	0,000,029	number_of_other	0,013,854	originate_from_its	0,000,012
longer_to_produce	0,000,047	number_of_systems	0,000,875	other_points_of	0,001,410
los_alamos_national	0,000,407	numbers_might_be	0,000,117	other_things_because	0,000,052
losers_are_the	0,000,013	nvdiia_tried_to	0,000,001	other_top_u	0,000,001
lost_track_by	0,000,001	oak_ridge_national	0,000,324	out_of_its	0,015,274
low_and_ultra	0,000,002	occupy_the_remaining	0,000,019	out_that_makes	0,000,005
lv_and_ulv	0,000,005	of_a_birds	0,000,006	out_the_top	0,000,347
machine_at_nasa	0,000,002	of_a_feather	0,001,372	out_what_you	0,002,308
made_changes_to	0,000,579	of_a_million	0,005,602	over_a_quarter	0,000,696
made_it_into	0,001,289	of_all_systems	0,000,340	own_workstations_with	0,000,001
mainly_at_the	0,000,332	of_clock_speed	0,000,019	owned_by_their	0,000,228
mainly_derived_from	0,000,208	of_ibm_power	0,000,002	percent_are_now	0,000,002
maintained_its_position	0,000,108	of_interest_include	0,000,366	percent_compared_to	0,000,128
make_the_large	0,000,059	of_its_hat	0,000,003	percent_of_the	0,042,833
makes_it_four	0,000,013	of_other_developments	0,000,019	percent_of_top	0,000,011
makes_the_battery	0,000,005	of_percent_the	0,000,024	percent_six_months	0,000,002
manufacturer_has_produced	0,000,001	of_performance_just	0,000,001	percent_the_company	0,000,001
manufacturers_and_their	0,000,061	of_service_so	0,000,037	performance_just_to	0,000,002
many_other_recent	0,000,027	of_supercomputers_in	0,000,008	performance_level_of	0,000,083
mark_on_the	0,003,382	of_systems_to	0,000,284	performance_of_all	0,000,800
market_as_a	0,000,489	of_the_best	0,062,027	performance_of_the	0,018,307
market_share_by	0,000,138	of_the_company	0,038,835	place_is_the	0,001,458
market_share_is	0,000,152	of_the_competition	0,001,160	plants_need_some	0,000,007
market_share_of	0,000,434	of_the_current	0,037,386	points_of_interest	0,001,950
matter_so_much	0,000,774	of_the_desktop	0,001,614	position_atop_the	0,000,024
maybe_with_a	0,000,115	of_the_german	0,024,021	position_in_the	0,022,546
mean_a_lot	0,000,412	of_the_hard	0,005,252	position_with_a	0,000,837
megawatt_mw_of	0,000,003	of_the_ivy	0,000,453	power_consumption_of	0,000,344
merged_with_the	0,002,068	of_the_low	0,009,776	power_efficiency_is	0,000,010
mhz_of_clock	0,000,003	of_the_most	0,268,014	powerful_as_its	0,000,010
might_also_stem	0,000,002	of_the_opengl	0,000,233	process_with_three	0,000,034
might_be_right	0,001,057	of_the_systems	0,003,476	processors_for_some	0,000,011
mind_on_short	0,000,003	of_the_top	0,009,629	processors_for_the	0,000,051
mistake_to_make	0,000,194	of_the_twice	0,000,337	processors_of_the	0,000,175
months_ago_and	0,000,356	of_the_university	0,024,486	processors_or_other	0,000,014
more_efficient_than	0,003,555	of_the_world	0,401,486	processors_will_feature	0,000,002
more_or_less	0,107,419	of_them_have	0,012,690	produce_enough_of	0,000,025
more_than_half	0,026,203	of_these_use	0,000,114	produced_by_now	0,000,002
more_than_megawatt	0,000,002	of_this_with	0,001,021	production_process_with	0,000,013
more_than_apple	0,000,001	of_top_systems	0,000,002	published_on_top	0,000,001
more_weeks_to	0,000,051	off_from_all	0,002,094	published_six_months	0,000,012
most_efficient_systems	0,000,002	on_a_giant	0,000,218	publishing_the_list	0,000,018
most_energy_efficient	0,000,073	on_a_windows	0,006,909	put_it_mildly	0,000,653

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

quarter_of_a	0,021,702	some_extra_time	0,000,142	the_company_would	0,001,418
quickly_even_if	0,000,024	some_of_them	0,093,847	the_current_version	0,004,499
quickly_pull_an	0,000,003	something_to_smile	0,000,048	the_damper_the	0,000,001
quiet_smile_and	0,000,039	somewhere_between_and	0,000,606	the_data_behind	0,000,052
rankings_did_not	0,000,001	spot_in_market	0,000,001	the_desktop_graphics	0,000,004
received_because_of	0,000,041	standard_for_mobile	0,000,037	the_discussion_at	0,000,313
recent_very_large	0,000,004	standard_linpack_benchmark	0,000,001	the_dual_core	0,000,098
reduces_costs_and	0,000,039	start_selling_the	0,000,009	the_end_of	0,510,039
reflects_will_be	0,000,001	starts_a_second	0,000,026	the_entry_level	0,000,328
refrained_from_sending	0,000,032	stem_from_the	0,001,284	the_expense_of	0,035,461
rejoice_over_a	0,000,051	still_hasn_t	0,000,369	the_failed_attempt	0,000,081
remain_unchanged_from	0,000,050	still_the_most	0,001,621	the_first_available	0,000,940
reports_the_highest	0,000,004	structures_that_the	0,000,111	the_first_ever	0,000,643
research_center_no	0,000,002	supercomputers_on_the	0,000,005	the_first_supercomputer	0,000,013
research_scientific_computing	0,000,006	supercomputing_conference_in	0,000,001	the_first_time	0,207,095
restructuring_the_company	0,000,009	supposed_to_originate	0,000,039	the_flood_in	0,000,349
resulted_in_a	0,011,494	supposed_to_reach	0,000,098	the_fourth_quarter	0,000,595
ridge_national_laboratory	0,000,298	supposed_to_start	0,000,139	the_german_original	0,000,134
round_out_the	0,000,540	surely_doesn_t	0,000,031	the_gk_of	0,000,001
rumors_about_a	0,000,061	system_at_no	0,000,015	the_gpu_is	0,000,014
rumors_about_the	0,000,215	system_is_a	0,004,918	the_hard_disk	0,006,162
running_the_standard	0,000,047	system_is_one	0,000,739	the_head_of	0,161,178
s_ames_research	0,000,026	system_on_the	0,003,019	the_highest_total	0,000,128
s_department_for	0,000,034	system_that_was	0,001,059	the_it_market	0,000,013
s_going_to	0,041,152	system_took_the	0,000,009	the_k_computer	0,000,004
s_mark_on	0,000,112	systems_account_for	0,000,012	the_large_notebook	0,000,001
s_most_powerful	0,000,835	systems_and_has	0,000,100	the_largest_share	0,000,666
s_not_the	0,014,578	systems_at_no	0,000,002	the_largest_u	0,000,148
s_one_year	0,000,061	systems_has_grown	0,000,022	the_last_minute	0,004,466
s_six_months	0,000,122	systems_in_china	0,000,029	the_last_system	0,000,171
s_still_unclear	0,000,013	systems_on_the	0,001,800	the_latest_list	0,000,109
s_to_stay	0,000,086	systems_showed_no	0,000,002	the_lead_of	0,003,547
sales_went_down	0,000,009	systems_to_and	0,000,008	the_level_to	0,000,482
sc_supercomputing_conference	0,000,001	systems_use_processors	0,000,003	the_linpack_benchmark	0,000,025
scheduled_for_this	0,000,063	t_be_affected	0,000,157	the_list_are	0,000,986
scientific_computing_center	0,000,005	t_have_any	0,027,544	the_list_back	0,000,041
second_place_is	0,000,096	t_made_it	0,000,345	the_losers_are	0,000,060
sector_is_growing	0,000,012	t_matter_so	0,000,339	the_low_and	0,001,597
seem_a_fatal	0,000,002	t_suit_them	0,000,057	the_mantra_is	0,000,346
seem_to_be	0,081,455	takes_longer_to	0,000,825	the_market_share	0,000,377
sell_off_their	0,000,034	taking_into_account	0,006,591	the_most_efficient	0,006,636
sending_us_a	0,000,087	talk_about_them	0,001,522	the_most_energy	0,000,214
service_so_that	0,000,383	th_edition_of	0,000,555	the_most_used	0,001,116
session_to_be	0,000,273	than_gigabit_ethernet	0,000,009	the_much_discussed	0,000,282
share_of_ibm	0,000,006	than_half_a	0,007,206	the_national_energy	0,000,108
share_of_percent	0,000,031	thanks_to_a	0,001,982	the_newest_edition	0,000,034
short_notice_and	0,000,105	thanks_to_the	0,018,735	the_newest_list	0,000,004
shortage_caused_by	0,000,026	that_apple_has	0,000,059	the_next_top	0,000,020
should_be_able	0,031,916	that_dominates_the	0,000,302	the_numbers_might	0,000,053
should_make_the	0,003,119	that_his_printing	0,000,003	the_oak_ridge	0,000,083
simply_didn_t	0,000,794	that_intel_is	0,000,035	the_old_opengl	0,000,001
since_we_began	0,000,235	that_is_to	0,096,356	the_one_or	0,004,928
six_months_ago	0,002,635	that_makes_it	0,008,567	the_ones_for	0,000,343
six_or_more	0,000,828	that_the_former	0,006,549	the_opengl_es	0,000,012
slightly_ahead_of	0,000,188	that_the_top	0,001,304	the_partially_built	0,000,017
slightly_down_from	0,000,001	that_was_at	0,003,602	the_previous_top	0,000,042
slip_us_a	0,000,005	that_was_the	0,072,744	the_processors_for	0,000,017
small_town_of	0,001,021	the_amd_opteron	0,000,044	the_remaining_space	0,000,511
smaller_competitors_with	0,000,001	the_battery_last	0,000,007	the_sales_went	0,000,003
smile_about_when	0,000,005	the_beginning_of	0,183,895	the_sc_supercomputing	0,000,001
smile_and_is	0,000,013	the_best_of	0,081,888	the_share_of	0,003,320
so_much_with	0,002,457	the_best_silicon	0,000,001	the_shop_counters	0,000,003
so_that_his	0,011,303	the_brand_new	0,000,398	the_silent_movie	0,000,027
solve_a_pesky	0,000,001	the_company_from	0,000,604	the_small_town	0,001,425

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

the_smaller_competitors	0,000,001	to_ibm_with	0,000,003	was_at_the	0,050,903
the_soc_market	0,000,001	to_its_performance	0,000,169	was_listed_at	0,000,043
the_southern_english	0,000,017	to_its_widespread	0,000,088	was_published_six	0,000,006
the_standard_for	0,002,701	to_make_the	0,144,608	was_the_first	0,092,970
the_standard_linpack	0,000,004	to_many_other	0,002,280	we_are_eagerly	0,000,017
the_system_is	0,018,661	to_occupy_the	0,005,525	we_began_publishing	0,000,017
the_systems_use	0,000,035	to_originate_from	0,000,545	weeks_to_sell	0,000,005
the_th_edition	0,000,486	to_produce_enough	0,000,235	well_advised_to	0,000,989
the_th_spot	0,000,007	to_provide_the	0,019,302	well_ahead_of	0,000,740
the_top_list	0,000,146	to_put_it	0,023,885	went_down_by	0,000,316
the_top_rankings	0,000,012	to_quickly_pull	0,000,014	what_a_coincidence	0,000,271
the_top_spot	0,000,138	to_sell_off	0,000,332	what_s_going	0,015,063
the_top_supercomputers	0,000,006	to_slip_us	0,000,018	what_you_don	0,002,117
the_top_systems	0,000,009	to_smile_about	0,000,082	when_it_tried	0,000,166
the_topic_of	0,008,206	to_start_selling	0,000,062	when_the_next	0,003,183
the_trends_it	0,000,002	to_stay_atop	0,000,010	which_had_already	0,004,650
the_twice_yearly	0,000,033	to_talk_about	0,030,228	which_have_instead	0,000,004
the_two_chinese	0,000,105	to_the_current	0,020,555	which_resulted_in	0,004,410
the_university_of	0,076,580	to_the_first	0,058,181	which_will_have	0,001,739
the_whims_of	0,001,160	to_the_list	0,012,775	which_will_most	0,000,203
the_world_s	0,091,548	to_their_orders	0,000,234	while_everything_is	0,000,059
their_orders_in	0,000,088	to_use_more	0,001,973	while_the_head	0,000,405
their_respective_owners	0,003,454	took_note_of	0,001,400	who_will_lead	0,000,414
them_have_made	0,000,117	took_the_top	0,000,154	widespread_use_at	0,000,025
there_are_rumors	0,000,288	top_list_by	0,000,002	will_be_the	0,078,372
they_are_allowed	0,002,130	top_list_is	0,000,005	will_feature_a	0,000,067
they_are_now	0,010,633	top_list_of	0,000,066	will_have_its	0,002,461
they_have_a	0,034,556	top_list_the	0,000,003	will_lead_the	0,001,108
they_will_now	0,000,381	top_list_was	0,000,003	will_most_likely	0,004,512
things_are_not	0,007,479	topic_of_a	0,000,186	will_now_have	0,000,893
things_because_of	0,000,228	total_of_systems	0,000,001	will_only_just	0,000,047
this_architecture_for	0,000,026	total_power_consumption	0,000,051	with_a_market	0,000,302
this_branch_of	0,003,289	town_of_kings	0,000,006	with_a_partner	0,000,997
this_doesn_t	0,007,173	track_by_now	0,000,007	with_a_performance	0,000,308
this_is_slightly	0,000,284	trademarks_and_copyrights	0,000,026	with_a_powerful	0,002,593
this_is_the	0,379,779	transistors_that_is	0,000,001	with_a_quiet	0,003,180
this_page_are	0,000,274	translation_of_the	0,018,904	with_more_than	0,017,920
this_surely_doesn	0,000,002	tried_to_completely	0,000,005	with_old_sandy	0,000,001
this_with_a	0,004,516	tried_to_slip	0,000,364	with_shader_model	0,000,001
thus_it_takes	0,000,022	twice_as_much	0,005,140	with_six_or	0,000,439
time_since_we	0,000,860	two_chinese_systems	0,000,001	with_somewhere_between	0,000,010
time_this_year	0,000,168	ultra_low_voltage	0,000,048	with_the_gk	0,000,005
times_as_fast	0,000,534	under_the_lead	0,000,940	with_the_ivy	0,000,100
times_as_powerful	0,000,091	university_of_mannheim	0,000,055	with_the_monster	0,000,439
tiny_ones_and	0,000,001	university_of_tennessee	0,001,399	with_the_pc	0,000,294
to_a_full	0,004,923	unperturbed_by_this	0,000,017	with_three_dimensional	0,000,272
to_achieve_a	0,010,051	up_from_six	0,000,008	working_on_a	0,006,684
to_an_internal	0,001,457	up_to_the	0,200,271	world_s_most	0,003,359
to_and_is	0,000,634	us_a_sample	0,000,065	would_be_more	0,021,234
to_be_happy	0,012,893	us_off_from	0,000,711	would_be_well	0,010,090
to_be_held	0,015,341	us_something_to	0,000,678	would_seem_a	0,000,813
to_be_raw	0,000,052	use_more_than	0,002,322	would_surely_say	0,000,069
to_blame_for	0,006,273	use_processors_with	0,000,002	xe_at_the	0,000,003
to_break_the	0,020,491	using_intel_processors	0,000,007	years_of_service	0,001,836
to_claim_the	0,003,630	vendors_are_supposed	0,000,004	yet_due_to	0,000,061
to_completely_cut	0,000,023	very_large_systems	0,000,079	you_don_t	0,439,225
to_hp_with	0,000,003	wafers_wth_nm	0,000,001		

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

The first half of next phrase-checked (with **Camera** corpus) 4-gram wordlist contains unfamiliar 4-grams whereas the second half holds all familiar to **Camera** 4-grams with their corresponding occurrences:

Overlapped lines, i.e. lines common for both files: **408**

Unfamiliar lines, i.e. lines from 1st file not encountered in 2nd file: **968**

a_certain_ivy_overgrown
a_cray_xt_system
a_feather_session_to
a_giant_ceramic_module
a_graphics_unit_with
a_meager_annual_plus
a_million_wafers_with
a_partner_is_working
a_powerful_graphics_core
a_previously_published_message
a_second_restructuring_attempt
a_strongly_downgraded_kepler
a_system_took_the
a_tegra_all_amongst
a_top_system_is
a_user_of_hpc
a_whole_received_because
a_wolf_speaking_of
about_a_tegra_all
about_ivy_bridges_and
about_possible_future_workstation
about_when_it_tried
accelerators_up_from_six
according_to_jon_peddie
account_for_almost_twice
account_that_intel_is
adapteva_would_surely_say
advanced_institute_for_computational
advertising_new_ultrabooks_with
advised_to_quickly_pull
ago_and_average_power
ago_and_mflops_watt
agreement_for_z_mainboards
alamos_national_laboratory_no
all_amongst_other_things
all_communication_concerning_the
all_current_gpgpu_implementations
all_systems_has_grown
all_using_nvidia_gpus
already_arrived_in_shops
already_gave_us_something
already_working_on_nm
also_refrained_from_sending
also_reports_the_highest
although_the_top_rankings
altix_machine_at_nasa
amd_manufacturer_has_produced
amd_opteron_family_with
ames_research_center_no
amongst_other_things_because
an_ibm_system_that
an_internal_diagram_with
an_ipad_competitor_out
an_sgi_altix_machine
and_also_refrained_from
and_average_power_efficiency
and_be_four_times
and_copyrights_on_this

and_france_round_out
and_has_now_systems
and_horst_simon_of
and_is_mainly_derived
and_jack_dongarra_of
and_makes_the_battery
and_merrily_philosophized_about
and_mflops_watt_one
and_now_wants_chips
and_nvidia_launches_rumors
and_offhandedly_commented_on
and_tegra_hitches_by
and_tegra_hitches_links
and_the_japanese_tsunami
and_the_trends_it
and_their_clients_wait
and_two_use_ati
and_ulv_at_the
angry_like_a_wolf
annoyed_nvidia_s_charismatic
any_bit_drivers_for
appears_that_apple_has
apple_has_changed_its
apple_publicly_and_offhandedly
architecture_for_graphics_cards
are_all_using_nvidia
are_bluegene_q_with
are_confirmed_to_use
are_not_so_super
are_now_using_intel
are_rumors_about_the
are_the_smaller_competitors
as_accelerators_up_from
as_for_the_ivy
as_his_business_sector
at_los_alamos_no
at_mhz_of_clock
at_the_riken_advanced
atoms_with_a_powerful
atop_the_newest_edition
atop_top_list_by
attempt_and_nvidia_launches
attempt_with_touchpad_and
average_power_efficiency_is
awaiting_the_silent_movie
barebones_with_old_sandy
based_systems_account_for
be_found_in_superphones
be_raw_in_santa
because_a_previously_published
began_publishing_the_list
behind_the_shop_counters
best_silicon_dice_receive
bit_drivers_for_the
bits_are_nothing_to
blame_for_the_delays
blog_about_ivy_bridges
boss_even_more_that

boss_started_restructuring_the
branch_under_the_lead
brand_new_barebones_with
brand_new_nm_production
break_the_petaflop_s
bridge_chip_s_launch
bridge_delay_should_make
bridge_generation_before_they
bridge_processors_for_some
bridge_s_launch_date
bridge_stocks_at_reasonable
bridges_and_tegra_hitches
build_out_that_makes
built_k_computer_had
business_sector_is_growing
but_ibm_stays_well
but_it_could_solve
but_was_dethroned_when
but_why_should_nvidia
by_hans_meuer_of
by_infiniband_technology_with
by_marcel_sieslack_intel
by_total_performance_slightly
c_t_by_marcel
called_jaguar_and_installed
card_geforce_gtx_and
certain_ivy_overgrown_bridge
charismatic_boss_even_more
china_keeps_increasing_its
chinese_systems_at_no
chinese_tianhe_a_system
chip_s_launch_date
chip_tegra_can_currently
chips_with_hd_instead
claim_the_th_spot
clients_wait_in_favor
combined_performance_of_all
combined_processors_will_feature
comes_with_somewhere_between
commented_on_the_weaker
communication_concerning_the_new
company_from_the_small
company_sells_more_gpu
compared_to_hp_with
compared_to_ibm_with
compatible_with_shader_model
competitor_out_of_its
competitors_like_larrabee_and
competitors_with_empty_warehouses
compiled_by_hans_meuer
completely_cut_us_off
computational_science_aics_in
computer_also_reports_the
computer_fujitsu_captured_the
computer_had_taken_the
computer_it_achieved_an
computer_maintained_its_position
computer_tops_petaflop_s

- An example of English phrase-checker 'Graphein' r.3 utilizing the Camera r.15 corpus (powered by: 9,181,275 1-grams; 124,669,942 2-grams; 477,829,381 3-grams; 879,557,846 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

concerning_the_new_kepler
confirmed_to_use_more
confuse_the_one_or
confusion_with_the_ivy
consigned_to_history_after
consumption_of_a_top
consumption_of_supercomputers_in
consumption_of_supercomputers_top
contributors_stiller_published_on
contributors_stiller_translation_of
copyrights_on_this_page
core_comes_with_somewhere
core_processors_of_the
cores_and_more_cores
cores_compatible_with_shader
could_feature_a_strongly
could_solve_a_pesky
cpu_dice_on_a
cray_xe_at_los
cray_xe_at_the
cray_xt_system_called
current_gpgpu_implementations_or
current_ipad_powervr_core
damper_the_it_market
debut_in_the_gk
department_for_workstation_and
desktop_graphics_card_geforce
dethroned_when_the_next
dice_on_a_giant
dice_receive_these_energy
dimensional_finfet_transistors_that
dinosaur_vj_vyomesh_joshi
disclosure_agreement_for_z
disk_shortage_caused_by
does_highlight_a_number
does_not_utilize_graphics
dominates_the_soc_market
dongarra_of_the_university
down_from_six_months
down_the_valley_ultimately
dreaming_of_equipping_its
drivers_for_the_powervr
dual_core_processors_of
e_k_has_cores
eagerly_awaiting_the_silent
easily_leaves_its_competitors
edition_of_the_top
efficiency_is_mflops_watt
efficient_supercomputers_are_bluegene
efficient_than_all_current
energy_efficient_supercomputers_are
energy_efficient_systems_on
english_town_of_ivybridge
entry_level_of_performance
epiphany_e_k_has
equip_brand_new_barebones
equip_tegra_soc_with
equipping_its_own_workstations
erich_strohmaier_and_horst
es_interface_the_standard
even_if_mainly_at
even_intel_has_lost
even_more_that_apple
everything_is_sprouting_and

extra_time_this_year
failed_attempt_with_touchpad
favor_of_the_low
feather_session_to_be
feature_a_graphics_unit
feature_a_strongly_downgraded
finfet_transistors_that_is
first_ever_to_break
first_supercomputer_to_achieve
followed_by_infiniband_technology
followed_by_the_and
for_bits_are_nothing
for_computational_science_aics
for_graphics_cards_at
for_mobile_graphics_will
for_now_with_systems
for_the_ivy_bridge
for_the_powervr_graphics
for_workstation_and_thin
former_and_manufacturer_has
found_in_superphones_and
four_cpu_dice_on
france_round_out_the
from_all_communication_concerning
from_its_large_brother
from_mflops_watt_six
from_sending_us_a
from_the_damper_the
from_the_old_opengl
fujitsu_captured_the_no
full_build_out_that
future_workstation_tablets_with
geforce_core_comes_with
geforce_gtx_and_also
generation_before_they_are
geometry_shaders_and_is
german_original_in_c
gk_chip_see_picture
gk_of_the_desktop
globalfoundries_can_rejoice_over
gpgpu_implementations_or_competitors
gpu_is_supposed_to
gpu_should_be_able
gpus_as_accelerators_up
gpus_to_accelerate_computation
graphics_cards_at_the
graphics_core_codename_valleyview
graphics_of_the_current
graphics_processors_or_other
graphics_unit_with_kepler
graphics_will_only_just
group_can_be_merged
growing_quickly_even_if
gtx_and_also_refrained
had_taken_the_no
hans_meuer_of_the
hard_disk_shortage_caused
has_lost_track_by
has_stabilized_for_now
have_annoyed_nvidia_s
have_any_bit_drivers
have_instead_of_only
have_its_debut_in
have_to_equip_brand

hd_instead_of_hd
heavily_advertising_new_ultrabooks
highest_total_power_consumption
his_boss_started_restructuring
his_business_sector_is
his_printing_group_can
history_after_years_of
hitches_by_admin_created
horst_simon_of_nersc
hp_dinosaur_vj_vyomesh
hp_is_meanwhile_dreaming
hp_maybe_with_a
hp_s_department_for
hp_starts_a_second
hsun_huang_s_super
huang_s_super_chip
ibm_kept_its_lead
ibm_power_processors_has
ibm_stays_well_ahead
ibm_system_that_was
implementations_or_competitors_like
in_a_meager_annual
in_nm_structures_by
in_superphones_and_tablets
in_systems_and_has
in_the_previous_top
infiniband_based_systems_account
infiniband_technology_with_systems
installed_at_the_oak
installed_at_the_riken
instead_of_hd_graphics
instead_of_only_shader
institute_for_computational_science
intel_already_gave_us
intel_causes_confusion_with
intel_continues_to_provide
intel_has_lost_track
intel_is_heavily_advertising
intel_probably_took_note
intel_s_plants_need
interface_the_standard_for
interludes_behind_the_shop
internal_diagram_with_more
internal_system_interconnect_technology
into_account_that_intel
introduce_optional_geometry_shaders
introduced_this_architecture_for
ipad_competitor_out_of
is_compiled_by_hans
is_growing_quickly_even
is_heavily_advertising_new
is_meanwhile_dreaming_of
is_mflops_watt_unchanged
is_mflops_watt_up
is_slightly_down_from
is_sprouting_and_leafing
is_the_chinese_tianhe
is_the_first_supercomputer
it_achieved_an_impressive
it_and_the_trends
it_could_feature_a
it_market_as_a
it_might_also_stem
it_s_imagination_technologies

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

it_s_processor_epiphany
it_will_pack_four
its_competitors_qualcomm_percent
its_lead_in_systems
its_mind_on_short
its_number_of_systems
its_own_workstations_with
ivy_bridge_chip_s
ivy_bridge_delay_should
ivy_bridge_generation_before
ivy_bridge_s_launch
ivy_bridges_and_tegra
ivy_overgrown_bridge_in
ivybridge_still_hasn_t
jack_dongarra_of_the
jaguar_and_installed_at
japan_and_france_round
japan_s_k_computer
jen_hsun_huang_s
jim_zarafana_obviously_proud
joshi_is_being_consigned
just_introduce_optional_geometry
just_introduced_this_architecture
k_computer_also_reports
k_computer_fujitsu_captured
k_computer_had_taken
k_computer_is_also
k_computer_is_the
k_computer_it_achieved
k_computer_maintained_its
k_computer_tops_petaflop
k_has_cores_admittedly
keeps_increasing_its_number
kepler_graphics_card_geforce
kept_its_lead_in
kings_langley_near_london
langley_near_london_easily
large_notebook_manufacturers_happy
larrabee_and_knights_corner
latest_list_the_th
launches_rumors_about_a
lead_in_systems_and
lead_of_todd_bradley
lead_the_discussion_at
leave_out_what_you
leaves_its_competitors_qualcomm
less_reasonably_scheduled_dates
level_of_performance_just
level_of_petaflop_s
like_a_wolf_speaking
like_larrabee_and_knights
list_are_confirmed_to
list_back_in_that
list_by_admin_created
list_does_highlight_a
list_moved_up_to
list_remain_unchanged_from
list_the_th_edition
list_the_top_list
list_was_listed_at
list_was_published_six
listed_at_position_in
lists_press_release_links
london_easily_leaves_its

longer_to_produce_enough
losers_are_the_smaller
lost_track_by_now
low_voltage_variants_lv
lv_and_ulv_at
machine_at_nasa_s
maintained_its_position_atop
make_the_large_notebook
make_the_ultrabook_manufacturers
manufacturer_has_produced_by
manufacturers_and_their_clients
many_other_recent_very
marcel_sieslack_intel_causes
mark_on_the_linpack
market_share_by_total
market_share_is_concerned
maybe_with_a_partner
meager_annual_plus_of
meanwhile_dreaming_of_equipping
megawatt_mw_of_power
merged_with_the_pc
merrily_philosophized_about_possible
message_simply_didn_t
meurer_of_the_university
mflops_watt_one_year
mflops_watt_six_months
mflops_watt_up_from
million_wafers_with_nm
mind_on_short_notice
mobile_graphics_will_only
mobile_ulp_geforce_core
models_and_merrily_philosophized
monster_processor_power_from
months_ago_and_average
months_ago_and_mflops
more_than_megawatt_mw
more_than_apple_publicly
more_weeks_to_sell
most_efficient_systems_on
most_energy_efficient_supercomputers
most_energy_efficient_systems
most_used_internal_system
movie_interludes_behind_the
much_discussed_cuda_operations
much_with_the_monster
must_have_annoyed_nvidia
mw_six_months_ago
near_london_easily_leaves
nersc_lawrence_berkeley_national
new_barebones_with_old
new_kepler_graphics_card
new_models_and_merrily
new_nm_production_process
new_ultrabooks_with_ivy
newcomer_adapteva_would_surely
newest_list_does_highlight
newest_list_was_listed
next_top_list_was
nine_systems_use_gpus
nm_atoms_with_a
nm_production_process_with
nm_structures_by_globalfoundries
nm_structures_that_the
no_other_company_sells

non_disclosure_agreement_for
not_so_super_as
not_the_brand_new
not_utilize_graphics_processors
note_of_this_with
notice_and_now_wants
now_clearly_the_no
now_have_to_equip
now_tracks_actual_power
now_using_intel_processors
now_wants_chips_with
numbers_might_be_right
nvidia_equip_tegra_soc
nvidia_gpus_to_accelerate
nvidia_launches_rumors_about
nvidia_only_just_introduced
nvidia_s_charismatic_boss
nvidia_tried_to_completely
obviously_proud_presented_new
of_a_million_wafers
of_a_top_system
of_all_systems_has
of_big_intel_clients
of_clock_speed_while
of_equipping_its_own
of_hp_s_department
of_ibm_power_processors
of_ivybridge_still_hasn
of_kings_langley_near
of_nersc_lawrence_berkeley
of_only_shader_units
of_percent_the_company
of_performance_just_to
of_supercomputers_in_a
of_supercomputers_top_now
of_systems_to_and
of_the_best_silicon
of_the_current_atoms
of_the_desktop_graphics
of_the_ivy_bridge
of_the_opengl_es
of_the_tegra_in
of_these_use_nvidia
off_their_sandy_bridge
offhandedly_commented_on_the
old_sandy_bridge_processors
on_a_giant_ceramic
on_a_windows_tablet
on_nm_atoms_with
on_the_latest_list
on_the_newest_list
on_the_weaker_graphics
on_top_supercomputing_sites
one_or_other_client
ones_and_is_supposed
ones_for_bits_are
only_just_introduce_optional
only_just_introduced_this
opengl_es_interface_the
opteron_family_with_systems
optional_geometry_shaders_and
or_competitors_like_larrabee
or_less_reasonably_scheduled
or_quadtrillion_calculations_per

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

org_about_ivy_bridges
org_all_trademarks_and
org_blog_about_ivy
org_contributors_stiller_published
org_contributors_stiller_translation
org_copyright_c_top
org_japan_s_k
org_lists_press_release
org_tags_whispers_copyright
original_in_c_t
originate_from_its_large
other_company_sells_more
other_recent_very_large
overall_ivy_bridge_delay
overgrown_bridge_in_the
own_workstations_with_that
pack_four_cpu_dice
page_are_owned_by
partially_built_k_computer
partner_is_working_on
pc_branch_under_the
percent_are_now_using
percent_compared_to_hp
percent_of_top_systems
percent_six_months_ago
percent_the_company_from
performance_just_to_claim
performance_level_of_petaflop
performance_of_the_tegra
performance_running_the_standard
performance_slightly_ahead_of
petaflop_s_on_the
petaflop_s_one_year
petaflop_s_performance_running
petaflop_s_six_months
petaflop_s_than_gigabit
petaflop_s_to_stay
philosophized_about_possible_future
place_is_the_chinese
plants_need_some_extra
position_atop_the_newest
position_with_a_performance
possible_future_workstation_tablets
power_consumption_of_supercomputers
power_efficiency_is_mflops
power_processors_has_stabilized
powerful_as_its_nearest
powerful_graphics_core_codename
powervr_graphics_of_the
presented_new_models_and
previous_top_just_six
previously_published_message_simply
printing_group_can_be
probably_took_note_of
process_with_three_dimensional
processor_epiphany_e_k
processor_power_from_ibm
processors_for_some_weeks
processors_for_the_largest
processors_has_stabilized_for
processors_of_the_ivy
processors_or_other_accelerators
processors_with_six_or
produced_in_nm_structures

production_process_with_three
proud_presented_new_models
provide_the_processors_for
publicly_and_offhandedly_commented
published_message_simply_didn
published_on_top_supercomputing
publishing_the_list_back
pull_an_ipad_competitor
quadrillion_calculations_per_second
quickly_even_if_mainly
quickly_pull_an_ipad
quiet_smile_and_is
rankings_did_not_change
raw_in_santa_clara
receive_these_energy_efficiency
recent_very_large_systems
reduces_costs_and_makes
reflects_will_be_the
refrained_from_sending_us
rejoice_over_a_quarter
remain_unchanged_from_june
reports_the_highest_total
restructuring_attempt_and_nvidia
resulted_in_a_meager
riken_advanced_institute_for
rumors_about_a_tegra
rumors_about_the_tegra
running_the_standard_linpack
s_charismatic_boss_even
s_department_for_workstation
s_going_to_occupy
s_k_computer_maintained
s_k_computer_tops
s_not_the_brand
s_on_the_linpack
s_one_year_ago
s_performance_running_the
s_plants_need_some
s_processor_epiphany_e
s_super_chip_tegra
s_than_gigabit_ethernet
s_to_stay_atop
said_top_editor_erich
sales_went_down_by
sandy_bridge_processors_for
sandy_bridge_stocks_at
scheduled_for_this_summer
science_aics_in_kobe
scientific_computing_center_no
second_restructuring_attempt_and
sector_is_growing_quickly
seem_a_fatal_mistake
seem_to_be_raw
sell_off_their_sandy
selling_the_dual_core
sells_more_gpu_designs
sending_us_a_sample
sgi_altix_machine_at
shaders_and_is_mainly
share_by_total_performance
share_of_ibm_power
share_of_percent_the
short_notice_and_now
should_make_the_large

should_nvidia Equip_tegra
sieslack_intel_causes_confusion
silent_movie_interludes_behind
silicon_dice_receive_these
simon_of_nersc_lawrence
since_we_began_publishing
six_months_ago_systems
slightly_ahead_of_cray
slightly_down_from_six
slip_us_a_non
small_town_of_kings
smaller_competitors_with_empty
smile_and_is_already
so_super_as_far
so_that_his_printing
socs_with_cores_compatible
solve_a_pesky_problem
some_extra_time_this
somewhere_between_and_shaders
southern_english_town_of
spot_in_market_share
stabilized_for_now_with
standard_for_mobile_graphics
standard_linpack_benchmark_application
start_selling_the_dual
started_restructuring_the_company
starts_a_second_restructuring
stay_atop_top_list
stays_well_ahead_of
stem_from_the_damper
still_in_second_place
stiller_published_on_top
stiller_translation_of_the
stocks_at_reasonable_prices
strohmaier_and_horst_simon
strongly_downgraded_kepler_architecture
structures_that_the_former
super_as_far_as
super_chip_tegra_can
supercomputer_to_achieve_a
supercomputers_are_bluegene_q
supercomputers_in_a_consistent
supercomputers_on_the_latest
supercomputers_top_now_tracks
supercomputing_conference_in_seattle
supposed_to_start_selling
surely_doesn't_matter
system_called_jaguar_and
system_interconnect_technology_systems
system_is_a_cray
system_on_the_newest
system_took_the_top
systems_account_for_almost
systems_showed_no_turnover
systems_to_and_is
systems_use_gpus_as
t_be_affected_though
t_by_marcel_sieslack
t_have_any_bit
tags_whispers_copyright_c
tegra_all_amongst_other
tegra_can_currently_be
tegra_combined_processors_will
tegra_gpu_should_be

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

tegra_hitches_by_admin
tegra_in_comparison_to
tegra_soccs_with_cores
teraflop_s_mark_on
teraflop_s_six_months
teraflops_at_mhz_of
than_all_current_gpgpu
than_gigabit_ethernet_ones
than_megawatt_mw_of
thanks_to_the_k
that_apple_has_changed
that_apple_publicly_and
that_dominates_the_soc
that_his_printing_group
that_intel_is_heavily
that_the_former_aml
that_the_top_systems
the_battery_last_longer
the_best_silicon_dice
the_brand_new_nm
the_chinese_tianhe_a
the_current_ipad_powervr
the_damper_the_it
the_desktop_graphics_card
the_discussion_at_sc
the_failed_attempt_with
the_first_available_pictures
the_first_supercomputer_to
the_flood_in_thailand
the_former_aml_manufacturer
the_gk_chip_see
the_gk_of_the
the_gpu_is_supposed
the_hard_disk_shortage
the_head_of_hp
the_highest_total_power
the_hp_dinosaur_vj
the_it_market_as
the_ivy_bridge_chip
the_ivy_bridge_generation
the_ivy_bridge_s
the_k_computer_also
the_k_computer_fujitsu
the_k_computer_is
the_k_computer_it
the_large_notebook_manufacturers
the_latest_list_the
the_lead_of_todd
the_linpack_benchmark_using
the_list_are_confirmed
the_list_moved_up
the_low_and_ultra
the_mobile_ulp_geforce
the_monster_processor_power
the_most_used_internal
the_much_discussed_cuda
the_new_kepler_graphics
the_newest_list_does
the_newest_list_was
the_next_top_list
the_ones_for_bits
the_opengl_es_interface
the_overall_ivy_bridge
the_partially_built_k

the_pc_branch_under
the_petaflop_s_barrier
the_powervr_graphics_of
the_previous_top_just
the_riken_advanced_institute
the_sales_went_down
the_share_of_ibm
the_silent_movie_interludes
the_smaller_competitors_with
the_southern_english_town
the_systems_use_processors
the_tegra_gpu_should
the_tegra_in_comparison
the_th_spot_increases
the_tianhe_a_system
the_top_rankings_did
the_top_supercomputers_on
the_top_systems_showed
the_trends_it_reflects
the_twice_yearly_list
the_two_chinese_systems
the_ultrabook_manufacturers_and
the_unified_shaders_of
the_upcoming_tegra_combined
the_weaker_graphics_performance
the_whims_of_big
their_clients_wait_in
their_orders_in_favor
their_sandy_bridge_stocks
them_have_made_changes
these_energy_efficiency_classifications
these_use_nvidia_chips
things_because_of_apple
thirty_nine_systems_use
this_architecture_for_graphics
this_is_slightly_down
this_page_are_owned
three_dimensional_finfet_transistors
tianhe_a_system_took
tianhe_a_system_with
tiny_ones_and_is
to_a_full_build
to_an_internal_diagram
to_break_the_petaflop
to_completely_cut_us
to_equip_brand_new
to_history_after_years
to_hp_with_systems
to_ibm_with_systems
to_jon_peddie_research
to_make_the_ultrabook
to_provide_the_processors
to_stay_atop_top
to_the_current_ipad
to_the_k_computer
to_the_list_moved
top_editor_erich_strohmaier
top_just_six_months
top_list_by_admin
top_list_is_compiled
top_list_the_top
top_list_was_published
top_now_tracks_actual
top_rankings_did_not

top_supercomputers_on_the
top_supercomputing_sites_http
top_systems_showed_no
topic_of_a_birds
tops_petaflop_s_to
total_combined_performance_of
total_performance_slightly_ahead
town_of_ivybridge_still
town_of_kings_langley
tracks_actual_power_consumption
transistors_that_is_to
trends_it_reflects_will
tried_to_completely_cut
twice_yearly_list_remain
two_chinese_systems_at
two_use_ati_radeon
two_use_cell_processors
ulp_geforce_core_comes
ultra_low_voltage_variants
ultrabook_manufacturers_and_their
ultrabooks_with_ivy_bridge
ulv_at_the_last
unified_shaders_of_model
unit_with_kepler_architecture
up_from_mflops_watt
up_from_six_months
upcoming_tegra_combined_processors
us_a_non_disclosure
us_something_to_smile
use_at_industrial_customers
use_gpus_as_accelerators
use_more_than_megawatt
use_processors_with_six
used_internal_system_interconnect
using_nvidia_gpus_to
utilize_graphics_processors_or
variants_lv_and_ulv
vj_vyomesh_joshi_is
voltage_variants_lv_and
vyomesh_joshi_is_being
wafers_with_nm_structures
wait_in_favor_of
wants_chips_with_hd
was_dethroned_when_the
was_listed_at_position
was_published_six_months
watt_one_year_ago
watt_six_months_ago
watt_up_from_mflops
weaker_graphics_performance_of
weeks_to_sell_off
well_advised_to_quickly
well_ahead_of_both
when_the_next_top
which_the_tegra_gpu
while_everything_is_sprouting
whims_of_big_intel
whispers_copyright_c_top
whole_received_because_of
why_should_nvidia_equip
widespread_use_at_industrial
will_have_its_debut
will_only_just_introduce
will_pack_four_cpu

- An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams; [124,669,942](#) 2-grams; [477,829,381](#) 3-grams; [879,557,846](#) 4-grams);
- The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;
- Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

with_a_powerful_graphics
with_cores_compatible_with
with_hd_instead_of
with_ivy_bridge_processors
with_nm_structures_that
with_old_sandy_bridge
with_that_knightly_power
with_the_gk_chip
with_the_ivy_bridge
with_the_monster_processor
with_the_pc_branch
with_three_dimensional_finfet
with_touchpad_and_webos
wolf_speaking_of_powervr
working_on_nm_atoms
workstation_and_thin_clients
workstation_tablets_with_us
workstations_with_that_knightly
would_seem_a_fatal
xe_at_los_alamos
xe_at_the_national
xt_system_called_jaguar
yearly_list_remain_unchanged
years_of_service_so
zarafana_obviously_proud_presented
zarafana_won_t_be

a_dozen_more_or	0,000,026	before_they_are_allowed	0,000,124	had_already_arrived_in	0,000,061
a_fatal_mistake_to	0,000,062	behind_it_and_the	0,000,096	half_a_dozen_more	0,000,515
a_few_more_weeks	0,000,202	being_consigned_to_history	0,000,001	has_changed_its_mind	0,000,023
a_full_build_out	0,000,001	best_of_the_best	0,000,442	has_produced_by_now	0,000,002
a_market_share_of	0,000,053	birds_of_a_feather	0,000,690	hasn_t_made_it	0,000,040
a_non_disclosure_agreement	0,000,055	branch_of_the_company	0,000,025	have_a_few_more	0,000,450
a_number_of_other	0,010,853	bridge_in_the_southern	0,000,001	have_made_changes_to	0,000,211
a_performance_level_of	0,000,006	but_only_the_best	0,000,049	highlight_a_number_of	0,000,035
a_quarter_of_a	0,018,585	but_the_whims_of	0,000,015	if_mainly_at_the	0,000,005
a_quiet_smile_and	0,000,022	but_things_are_not	0,000,100	in_a_consistent_fashion	0,000,089
a_total_of_systems	0,000,001	by_the_and_opteron	0,000,003	in_comparison_to_the	0,002,669
about_the_top_list	0,000,004	by_the_flood_in	0,000,018	in_favor_of_apple	0,000,001
according_to_an_internal	0,000,033	by_their_respective_owners	0,000,144	in_favor_of_the	0,014,027
according_to_the_first	0,001,077	can_be_merged_with	0,000,079	in_market_share_by	0,000,001
achieve_a_performance_level	0,000,002	can_currently_be_found	0,000,012	in_second_place_is	0,000,003
actual_power_consumption_of	0,000,001	can_rejoice_over_a	0,000,001	in_that_the_top	0,000,004
after_the_failed_attempt	0,000,009	cards_at_the_end	0,000,009	in_the_fourth_quarter	0,000,209
after_years_of_service	0,000,093	caused_by_the_flood	0,000,020	in_the_latest_list	0,000,001
all_trademarks_and_copyrights	0,000,001	causes_confusion_with_the	0,000,001	in_the_southern_english	0,000,003
allowed_to_talk_about	0,000,055	changed_its_mind_on	0,000,001	increasing_its_number_of	0,000,004
almost_twice_as_much	0,000,088	changes_to_their_orders	0,000,001	intel_doesn_t_have	0,000,001
also_one_of_the	0,003,765	company_would_be_well	0,000,003	is_already_working_on	0,000,048
also_stem_from_the	0,000,014	comparison_to_the_current	0,000,010	is_also_one_of	0,001,914
and_for_the_much	0,000,011	computer_is_also_one	0,000,002	is_being_consigned_to	0,000,006
and_installed_at_the	0,000,021	computer_is_the_first	0,000,006	is_mainly_derived_from	0,000,074
and_is_already_working	0,000,001	continues_to_provide_the	0,000,047	is_now_clearly_the	0,000,006
and_is_now_clearly	0,000,001	contrary_to_many_other	0,000,008	is_one_of_the	0,117,469
and_is_supposed_to	0,000,741	costs_and_makes_the	0,000,002	is_still_the_most	0,000,812
and_the_ones_for	0,000,009	currently_be_found_in	0,000,002	is_supposed_to_originate	0,000,011
and_thus_it_takes	0,000,006	cut_us_off_from	0,000,235	is_supposed_to_reach	0,000,013
and_ultra_low_voltage	0,000,007	data_behind_it_and	0,000,001	is_the_first_time	0,005,474
are_allowed_to_talk	0,000,027	delay_should_make_the	0,000,002	is_to_blame_for	0,000,648
are_eagerly_awaiting_the	0,000,005	derived_from_the_old	0,000,243	is_working_on_a	0,000,498
are_nothing_to_be	0,000,051	diagram_with_more_than	0,000,002	it_appears_that_apple	0,000,001
are_now_followed_by	0,000,009	didn_t_suit_them	0,000,018	it_could_solve_a	0,000,005
are_owned_by_their	0,000,120	doesn_t_have_any	0,001,909	it_does_not_utilize	0,000,013
are_supposed_to_start	0,000,015	doesn_t_matter_so	0,000,123	it_four_times_as	0,000,004
as_a_whole_received	0,000,003	doesn_t_mean_a	0,000,351	it_into_the_spotlight	0,000,003
as_far_as_the	0,040,511	dozen_more_or_less	0,000,021	it_must_have_annoyed	0,000,014
as_fast_as_the	0,003,296	due_to_its_performance	0,000,003	it_reflects_will_be	0,000,001
as_its_nearest_competitor	0,000,001	due_to_its_widespread	0,000,023	it_s_still_unclear	0,000,012
as_powerful_as_its	0,000,009	edition_of_the_twice	0,000,004	it_takes_longer_to	0,000,337
as_the_current_version	0,000,028	efficient_systems_on_the	0,000,002	it_tried_to_slip	0,000,006
as_the_market_share	0,000,010	energy_research_scientific_computing	0,000,006	it_would_seem_a	0,000,264
at_los_alamos_national	0,000,075	ethernet_is_still_the	0,000,003	its_debut_in_the	0,000,019
at_nasa_s_ames	0,000,007	ever_to_break_the	0,000,008	its_position_atop_the	0,000,002
at_position_in_the	0,000,172	expense_of_the_competition	0,000,002	its_widespread_use_at	0,000,003
at_the_beginning_of	0,062,221	far_as_the_market	0,000,038	just_six_months_ago	0,000,027
at_the_end_of	0,213,050	fast_as_the_current	0,000,005	just_to_claim_the	0,000,002
at_the_expense_of	0,026,232	fatal_mistake_to_make	0,000,002	last_system_on_the	0,000,001
at_the_last_minute	0,002,596	few_more_weeks_to	0,000,014	lawrence_berkeley_national_laboratory	0,000,257
at_the_national_energy	0,000,004	first_time_since_we	0,000,231	level_to_the_list	0,000,001
at_the_oak_ridge	0,000,018	for_almost_twice_as	0,000,002	list_is_compiled_by	0,000,002
at_the_sc_supercomputing	0,000,001	for_the_largest_share	0,000,031	list_of_the_world	0,000,060
average_power_consumption_of	0,000,015	for_the_much_discussed	0,000,005	los_alamos_national_laboratory	0,000,328
back_in_that_the	0,000,001	four_times_as_fast	0,000,073	low_and_ultra_low	0,000,002
be_able_to_manage	0,001,568	four_times_as_powerful	0,000,007	made_changes_to_their	0,000,012
be_four_times_as	0,000,069	from_six_months_ago	0,000,004	made_it_into_the	0,000,384
be_merged_with_the	0,000,208	from_the_small_town	0,000,031	mainly_at_the_expense	0,000,018
be_more_efficient_than	0,000,385	gave_us_something_to	0,000,043	mainly_derived_from_the	0,000,095
be_the_topic_of	0,000,120	gigabit_ethernet_is_still	0,000,003	makes_it_four_times	0,000,001
be_well_advised_to	0,000,551	going_to_occupy_the	0,000,012	makes_the_battery_last	0,000,001
because_it_saves_space	0,000,008	graphics_card_geforce_gtx	0,000,002	market_as_a_whole	0,000,182
because_of_the_hard	0,000,063	graphics_performance_of_the	0,000,002	market_share_of_percent	0,000,009

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayce.com

matter_so_much_with	0,000,003	s_most_powerful_supercomputers	0,000,008	the_standard_linpack_benchmark	0,000,001
mhz_of_clock_speed	0,000,002	s_six_months_ago	0,000,010	the_system_is_one	0,000,051
might_also_stem_from	0,000,002	sc_supercomputing_conference_in	0,000,001	the_th_edition_of	0,000,224
mistake_to_make_the	0,000,017	second_place_is_the	0,000,015	the_top_list_is	0,000,003
more_efficient_than_all	0,000,001	service_so_that_his	0,000,002	the_top_list_of	0,000,027
more_or_less_reasonably	0,000,003	session_to_be_held	0,000,004	the_top_list_the	0,000,003
more_than_half_a	0,005,059	shortage_caused_by_the	0,000,011	the_topic_of_a	0,000,141
moved_up_to_the	0,000,378	should_be_able_to	0,030,897	the_university_of_mannheim	0,000,007
nasa_s_ames_research	0,000,026	simply_didn_t_suit	0,000,001	the_university_of_tennessee	0,000,523
national_energy_research_scientific	0,000,008	six_months_ago_and	0,000,051	the_world_s_most	0,003,182
need_some_extra_time	0,000,002	six_or_more_cores	0,000,001	there_are_rumors_about	0,000,018
nerves_seem_to_be	0,000,006	smile_about_when_it	0,000,004	they_are_allowed_to	0,001,626
newest_edition_of_the	0,000,017	so_much_with_the	0,000,538	they_are_now_followed	0,000,005
nothing_to_be_happy	0,000,052	some_of_them_have	0,002,856	they_have_a_few	0,000,138
now_followed_by_the	0,000,032	something_to_smile_about	0,000,013	they_will_now_have	0,000,041
number_of_other_developments	0,000,005	still_hasn_t_made	0,000,012	things_are_not_so	0,000,454
number_of_systems_to	0,000,034	still_the_most_used	0,000,020	this_branch_of_the	0,001,398
oak_ridge_national_laboratory	0,000,293	supposed_to_originate_from	0,000,008	this_doesn_t_mean	0,001,385
occupy_the_remaining_space	0,000,003	system_is_one_of	0,000,294	this_is_the_first	0,014,352
of_a_birds_of	0,000,001	system_that_was_the	0,000,030	this_surely_doesn_t	0,000,002
of_a_feather_session	0,000,006	systems_and_has_now	0,000,001	this_with_a_quiet	0,000,009
of_service_so_that	0,000,016	systems_has_grown_to	0,000,003	thus_it_takes_longer	0,000,001
of_the_company_would	0,000,117	systems_on_the_list	0,000,003	time_since_we_began	0,000,004
of_the_german_original	0,000,031	systems_use_processors_with	0,000,001	times_as_fast_as	0,000,356
of_the_hard_disk	0,000,603	t_made_it_into	0,000,010	times_as_powerful_as	0,000,057
of_the_low_and	0,000,129	t_matter_so_much	0,000,271	to_achieve_a_performance	0,000,019
of_the_most_efficient	0,000,464	takes_longer_to_produce	0,000,007	to_and_is_now	0,000,016
of_the_most_energy	0,000,014	taking_into_account_that	0,000,288	to_be_happy_about	0,000,272
of_the_systems_use	0,000,005	th_edition_of_the	0,000,115	to_be_held_at	0,000,982
of_the_top_list	0,000,007	than_half_a_dozen	0,000,883	to_blame_for_the	0,001,162
of_the_twice_yearly	0,000,002	thanks_to_a_full	0,000,003	to_claim_the_th	0,000,008
of_the_university_of	0,012,200	that_is_to_blame	0,000,077	to_its_widespread_use	0,000,040
of_the_world_s	0,035,862	that_makes_it_four	0,000,003	to_many_other_recent	0,000,003
of_them_have_made	0,000,083	that_was_at_the	0,000,723	to_occupy_the_remaining	0,000,003
of_this_with_a	0,000,103	that_was_the_first	0,002,714	to_originate_from_its	0,000,005
off_from_all_communication	0,000,227	the_amd_opteron_family	0,000,001	to_produce_enough_of	0,000,002
on_short_notice_and	0,000,051	the_beginning_of_march	0,000,497	to_put_it_mildly	0,000,634
on_the_linpack_benchmark	0,000,002	the_best_of_the	0,009,191	to_quickly_pull_an	0,000,002
on_the_list_are	0,000,138	the_company_from_the	0,000,123	to_sell_off_their	0,000,010
on_this_page_are	0,000,180	the_company_would_be	0,000,263	to_smile_about_when	0,000,005
one_of_the_most	0,146,521	the_data_behind_it	0,000,002	to_start_selling_the	0,000,004
only_the_best_of	0,000,172	the_dual_core_processors	0,000,005	to_talk_about_them	0,000,561
orders_in_favor_of	0,000,004	the_end_of_march	0,001,549	to_the_first_available	0,000,134
other_points_of_interest	0,000,163	the_entry_level_of	0,000,025	to_their_orders_in	0,000,001
other_things_because_of	0,000,007	the_expense_of_the	0,010,099	to_use_more_than	0,000,764
out_of_its_hat	0,000,003	the_first_ever_to	0,000,049	took_note_of_this	0,000,075
out_that_makes_it	0,000,001	the_first_time_since	0,007,919	took_the_top_spot	0,000,002
out_what_you_don	0,000,022	the_german_original_in	0,000,006	top_list_of_the	0,000,009
over_a_quarter_of	0,000,593	the_last_system_on	0,000,001	total_power_consumption_of	0,000,010
owned_by_their_respective	0,000,014	the_level_to_the	0,000,034	trademarks_and_copyrights_on	0,000,003
percent_of_the_systems	0,000,007	the_list_back_in	0,000,003	translation_of_the_german	0,000,087
performance_of_all_systems	0,000,003	the_losers_are_the	0,000,007	tried_to_slip_us	0,000,001
points_of_interest_include	0,000,133	the_market_share_is	0,000,009	twice_as_much_performance	0,000,001
position_in_the_previous	0,000,009	the_most_efficient_systems	0,000,002	under_the_lead_of	0,000,909
power_consumption_of_a	0,000,052	the_most_energy_efficient	0,000,068	us_off_from_all	0,000,030
processors_will_feature_a	0,000,001	the_national_energy_research	0,000,009	vendors_are_supposed_to	0,000,004
produce_enough_of_them	0,000,001	the_newest_edition_of	0,000,024	was_at_the_beginning	0,000,768
published_six_months_ago	0,000,001	the_numbers_might_be	0,000,016	was_the_first_ever	0,000,067
quarter_of_a_million	0,001,934	the_oak_ridge_national	0,000,038	we_are_eagerly_awaiting	0,000,001
received_because_of_the	0,000,011	the_one_or_other	0,000,110	we_began_publishing_the	0,000,001
research_scientific_computing_center	0,000,004	the_processors_for_the	0,000,001	what_s_going_to	0,002,250
round_out_the_top	0,000,002	the_sc_supercomputing_conference	0,000,001	what_you_don_t	0,002,086
s_ames_research_center	0,000,026	the_small_town_of	0,000,780	when_it_tried_to	0,000,155
s_mark_on_the	0,000,017	the_standard_for_mobile	0,000,004	which_had_already_arrived	0,000,020

An example of English phrase-checker 'Graphein' r.3 utilizing the Gamera r.15 corpus (powered by: [9,181,275](#) 1-grams: [124,669,942](#) 2-grams: [477,829,381](#) 3-grams: [879,557,846](#) 4-grams):

The dump contains: **5** unfamiliar 1-grams out of **806**; **159** unfamiliar 2-grams out of **1526**; **600** unfamiliar 3-grams out of **1538**; **968** unfamiliar 4-grams out of **1376**;

Two top500.org articles were x-gram-checked; Downloaded from www.sanmayca.com

which_have_instead_of	0,000,003	will_lead_the_discussion	0,000,001	with_somewhere_between_and	0,000,004
which_resulted_in_a	0,000,638	will_now_have_to	0,000,201	won_t_be_affected	0,000,106
which_will_have_its	0,000,059	with_a_market_share	0,000,029	working_on_a_windows	0,000,033
which_will_most_likely	0,000,059	with_a_partner_is	0,000,026	world_s_most_powerful	0,000,157
while_the_head_of	0,000,090	with_a_performance_of	0,000,071	would_be_more_efficient	0,000,261
who_will_lead_the	0,000,096	with_a_quiet_smile	0,000,682	would_be_well_advised	0,000,392
will_be_the_topic	0,000,056	with_more_than_half	0,000,218	yet_due_to_its	0,000,003
will_feature_a_graphics	0,000,001	with_six_or_more	0,000,070	you_don_t_need	0,024,607

Enjoy!

Kaze,

2012-Apr-21

Japan's K Computer Tops 10 Petaflop/s to Stay Atop TOP500 List

By *admin*

Created 2011-11-11 11:11



BERKELEY, Calif.; KNOXVILLE, Tenn.; and MANNHEIM, Germany (Nov. 14, 2011)—Japan's "K Computer" maintained its position atop the newest edition of the TOP500 List of the world's most powerful supercomputers, thanks to a full build-out that makes it four times as powerful as its nearest competitor. Installed at the RIKEN Advanced Institute for Computational Science (AICS) in Kobe, Japan, the K Computer it achieved an impressive 10.51 Petaflop/s on the Linpack benchmark using 705,024 SPARC64 processing cores.

The K Computer is the first supercomputer to achieve a performance level of 10 Petaflop/s, or 10 quadrillion calculations per second. In June 2011, the partially built K computer had taken the No. 1 position with a performance of 8.16 Petaflop/s. Contrary to many other recent very large systems, it does not utilize graphics processors or other accelerators. The K Computer is also one of the most energy efficient systems on the list.

Still in second place is the Chinese Tianhe-1A system with 2.57 Petaflop/s performance. One year ago, the Tianhe-1A system took the top spot, but was dethroned when the next TOP500 list was published six months ago.

In fact, the Top 10 supercomputers on the latest list – the 38th edition of the twice-yearly list – remain unchanged from June 2011. The latest list, the data behind it and the trends it reflects will be the topic of a Birds-of-a-Feather session to be held at 5:30 p.m. Tuesday, Nov. 15, at the SC11 supercomputing conference in Seattle.

"This is the first time since we began publishing the list back in 1993 that the top 10 systems showed no turnover," said TOP500 editor Erich Strohmaier, who will lead the discussion at SC11.

The largest U.S. system is a Cray XT5 system called Jaguar and installed at the Oak Ridge National Laboratory, with a 1.75 Petaflop/s performance running the standard Linpack benchmark application. Other top U.S. systems include Cielo, a Cray XE6 at Los Alamos National Laboratory (No. 6); Pleiades, an SGI Altix machine at NASA's Ames Research Center (No.7); Hopper, a Cray XE6 at the National Energy Research Scientific Computing Center (No. 8); and Roadrunner, an IBM system that was the first ever to break the Petaflop/s barrier, at Los Alamos (No. 10). Systems in China, Japan and France round out the Top 10.

Although the top rankings did not change, the newest list does highlight a number of other developments. For example:

- China keeps increasing its number of systems to 75 and is now clearly the No. 2 country, as a user of HPC, ahead of Japan, UK, France, and Germany.
- The two Chinese systems at No.2 and No. 4 and the Japanese Tsubame 2.0 system at No. 5 are all using NVIDIA GPUs to accelerate computation.
- Thirty-nine systems use GPUs as accelerators (up from 17 six months ago), 35 of these use NVIDIA chips, two use Cell processors, and two use ATI Radeon.
- Already, 62 percent of the systems use processors with six or more cores.
- Intel continues to provide the processors for the largest share (76.8 percent) of TOP500 systems.
- Thanks to the K computer Fujitsu captured the No. 2 spot in market share by total performance slightly ahead of Cray, but IBM stays well ahead of both.

With every list, the entry level of performance just to claim the 500th spot increases. In the latest list, the level to the list moved up to the 50.9 Teraflop/s mark on the Linpack benchmark, compared to 39.1 Teraflop/s six months ago. The last system on the newest list was listed at position 305 in the previous TOP500 just six months ago. Total combined performance of all 500 systems has grown to 74.2 Petaflop/s, compared to 58.7 Petaflop/s six months ago and 43.7 Petaflop/s one year ago.

Other points of interest include:

- A total of 384 systems (76.8 percent) are now using Intel processors. This is slightly down from six months ago (386 systems 77.2 percent).
- They are now followed by the AMD Opteron family with 63 systems (12.6 percent), down from 66.
- The share of IBM Power processors has stabilized for now with 49 systems (9.8 percent), up from 45.
- Gigabit Ethernet is still the most-used internal system interconnect technology (223 systems, down from 230 systems), due to its widespread use at industrial customers, followed by InfiniBand technology with 213 systems, up from 208 systems. However, InfiniBand-based systems account for almost twice as much performance (28.7 Petaflop/s) than Gigabit Ethernet ones (14.2 Petaflop/s).
- IBM kept its lead in systems and has now 223 systems (44.6 percent) compared to HP with 140 systems (28.0 percent). HP had 146 systems (29.2 percent) six months ago, compared to IBM with 218 systems (43.6 percent).

Power consumption of supercomputers

TOP500 now tracks actual power consumption of supercomputers in a consistent fashion.

- 29 systems on the list are confirmed to use more than 1 megawatt (MW) of power.
- The No. 1 system, the K computer also reports the highest total power consumption of 12.66 MW. Yet due to its performance, the system is one of the most efficient systems on the list, delivering 830 Mflops/watt.
- The most energy efficient supercomputers are BlueGene/Q with 2,029 Mflops/watt.
- Average power efficiency is 282 Mflops/watt (up from 248 Mflops/watt six months ago and 219 Mflops/watt one year ago).
- Average power consumption of a TOP10 system is 4.56 MW (up from 4.3 MW six months ago) and average power efficiency is 464 Mflops/watt (unchanged).

About the TOP500 List

The TOP500 list is compiled by Hans Meuer of the University of Mannheim, Germany; Erich Strohmaier and Horst Simon of NERSC/Lawrence Berkeley National Laboratory; and Jack Dongarra of the University of Tennessee, Knoxville. For more information, visit www.top500.org [1]

Copyright (c) 2000-2011 TOP500.Org | All trademarks and copyrights on this page are owned by their respective owners.

Source URL: <http://www.top500.org/lists/2011/11/press-release>

Links:

[1] <http://www.top500.org>

About Ivy Bridges and Tegra Hitches

By *admin*

Created 2012-04-10 03:13



[Andreas Stiller](#) [1]

(Translation of the German original in c't by Marcel Sieslack)

Intel causes confusion with the Ivy Bridge's launch date, HP starts a second restructuring attempt and Nvidia launches rumors about a Tegra 4 – all (amongst other things) because of Apple.

While everything is sprouting and leafing, Intel's plants need some extra time this year. And so, a certain ivy-overgrown bridge in the southern English town of Ivybridge still hasn't made it into the spotlight. As for the Ivy Bridge chip's launch date, even Intel has lost track by now: According to an internal diagram with more than half a dozen more or less reasonably scheduled dates, vendors are supposed to start selling the dual-core processors of the Ivy Bridge generation before they are allowed to talk about them. We are eagerly awaiting the silent movie interludes behind the shop counters. In any case, Intel already gave us something to smile about when it tried to slip us a non-disclosure agreement for Z77 mainboards, which had already arrived in shops.

Apparently, it's not the brand new 22 nm production process with three dimensional FinFET transistors that is to blame for the delays, but the whims of big Intel clients. Allegedly, some of them have made changes to their orders in favor of the low and ultra-low voltage variants (LV and ULV) at the last minute. But only the best of the best silicon dice receive these energy efficiency classifications, and thus it takes longer to produce enough of them. Additionally, it appears that Apple has changed its mind on short notice and now wants chips with HD 4000 instead of HD 2500 graphics, which have 16 instead of only 6 shader units. However, taking into account that Intel is heavily advertising new ultrabooks with Ivy Bridge processors, it would seem a fatal mistake to make the ultrabook manufacturers and their clients wait in favor of Apple.

Meanwhile, the overall Ivy Bridge delay should make the large notebook manufacturers happy: they have a few more weeks to sell off their Sandy Bridge stocks at reasonable prices. The losers are the smaller competitors with empty warehouses. They will now have to equip brand new barebones with old Sandy Bridge processors for some weeks, which will most likely, to put it mildly, confuse the one or other client.

Down the Valley

Ultimately, it might also stem from the "damper" the IT market as a whole received because of the hard disk shortage caused by the flood in Thailand: In the fourth quarter, the sales went down by 5.9 percent, which resulted in a meager annual plus of 1.3 percent. Unperturbed by this, Globalfoundries can rejoice over a quarter of a million wafers with 32 nm structures that the former AMD manufacturer has produced by now. And yet, Intel probably took note of this with a quiet smile and is already working on 22 nm Atoms with a powerful graphics core – codename "Valleyview". The GPU is supposed to originate from its large brother, Ivy Bridge, and be four times as fast as the current version. This doesn't mean a lot, but it could solve a pesky problem: Intel doesn't have any 64-bit drivers for the PowerVR graphics of the current Atoms, and the ones for 32 bits are nothing to be happy about.

Angry Like a Wolf

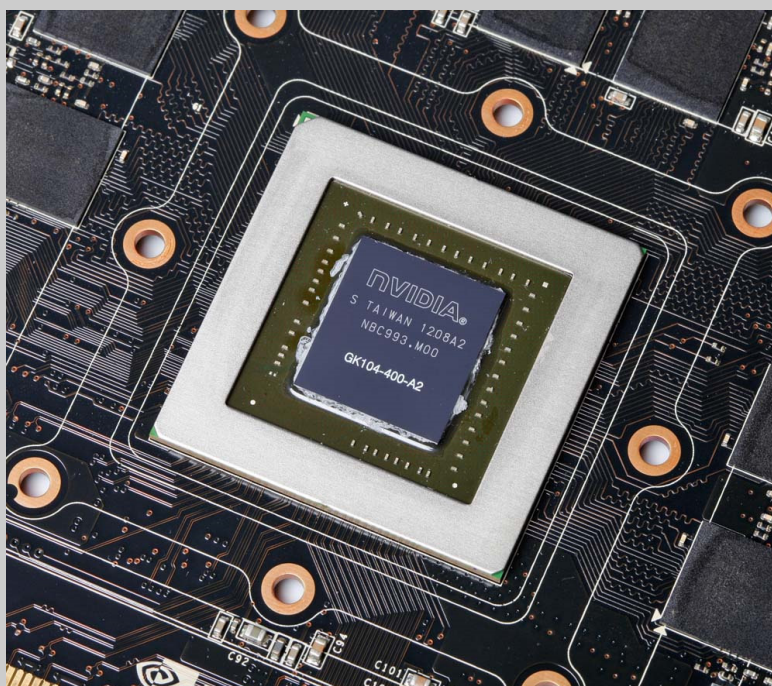
Speaking of PowerVR: According to Jon Peddie Research, it's Imagination Technologies, maker of PowerVR, that dominates the SoC market. Reportedly, no other company sells more GPU designs: With a market share of 50 percent the company from the small town of Kings Langley near London easily leaves its competitors Qualcomm (33 percent), Vivante (4.8 percent) and Nvidia (3.2 percent) behind. Jen-Hsun Huang's "super chip" Tegra can currently be found in "superphones" and tablets, but things are not so "super" as far

as the market share is concerned. Actually, regarding some topics, nerves seem to be raw in Santa Clara: For instance, Nvidia tried to completely cut us off from all communication concerning the new Kepler graphics card GeForce GTX 680 and also refrained from sending us a sample, because a previously published message simply didn't suit them.

It must have annoyed Nvidia's charismatic boss even more that Apple publicly and offhandedly commented on the weaker graphics performance of the Tegra 3 in comparison to the current iPad PowerVR core. That was at the beginning of March. And now, what a coincidence, there are rumors about the Tegra 4: It could feature a strongly downgraded Kepler architecture, which will have its debut in the GK104 of the desktop graphics card GeForce GTX 680. Supposedly, the mobile ULP GeForce core comes with somewhere between 32 and 64 shaders. Well, the numbers might be right, but why should Nvidia equip Tegra SoCs with cores compatible with shader model 5.1 when version 3.0, scheduled for this summer, of the OpenGL ES interface – the standard for mobile graphics – will only just introduce optional geometry shaders and is mainly derived from the old OpenGL 3? And for the much-discussed CUDA operations, which the Tegra 4 GPU should be able to manage, the unified shaders of model 4.0 are sufficient. For smartphone chips, the mantra is: Leave out what you don't need, because it saves space, reduces costs and makes the battery last longer.

This surely doesn't matter so much with the monster processor Power7+ from IBM. According to the first available pictures, it will pack four CPU dice on a giant ceramic module. It's still unclear, however, what's going to occupy the remaining space. Newcomer Adapteva would surely say: Cores, cores and more cores. Its processor Epiphany E4K has 4096 cores – admittedly, tiny ones – and is supposed to reach 5.632 teraflops at 700 MHz of clock speed while, nonetheless, consuming only 80 watts. If correct, the chip, produced in 28 nm structures by Globalfoundries, would be more efficient than all current GPGPU implementations or competitors like Larrabee and Knights Corner.

HP is meanwhile dreaming of equipping its own workstations with that knightly power. While the head of HP's department for workstation and thin clients, Jim Zarafana – obviously proud – presented new models and merrily philosophized about possible future workstation tablets with us, his boss started restructuring the company, again. Zarafana won't be affected though, as his business sector is growing quickly – even if mainly at the expense of the competition. However, the HP dinosaur "VJ" (Vyomesh Joshi) is being consigned to history after 31 years of service so that his Printing Group can be merged with the PC branch under the lead of Todd Bradley. After the failed attempt with Touchpad and WebOS, this branch of the company would be well advised to quickly pull an iPad competitor out of its hat. Allegedly, HP – maybe with a partner -- is working on a Windows 8 tablet.



According to rumors, the upcoming Tegra 4 combined processors will feature a graphics unit with Kepler architecture. Nvidia only just introduced this architecture for graphics cards at the end of March, with the GK104 chip (see picture).

- [whispers](#)

Copyright (c) 2000-2011 TOP500.Org | All trademarks and copyrights on this page are owned by their respective owners.

Source URL: http://www.top500.org/blog/2012/04/10/about_ivy_bridges_and_tegra_hitches

Links:

[1] <http://www.top500.org/contributors#stiller>

Project 'GAMERA' revision 15

(Semi-Powerful Console ([Windows](#) & [Linux](#)) Tools & gigabytes of English texts, downloadable from www.sanmayce.com)

WHERE THE WORD COUNTS: a FREE sub-linear English phrase-checker



Yoshi (Filelist creator and more, 32bit console application), revision 7-

Kazuya (Text decompressor-finder-dumper, 32bit console application), revision 19

GRAFFITH (Text decompressor-finder-dumper, 32bit console application), revision 2++_ **Graphain**

Leprechaun_x-leton (Fast and Greedy x-gram Ripper, 32bit/64bit Linux/Windows console application), revision 15fixfix

WinRAR archive in 32 624MB volumes • Required HDD/SSD space: 24.4 GB (*ready to go when extracted on D:*) • 2012_APR_20

Gamera r.15 x-gram corpora stats: 1/2/3/4/8 respectively with 009,181,275/124,669,942/477,829,381/879,557,846/812,576,024 x-grams. Gamera rips in sub-linear time (size of texts being checked matters a bit) all x-grams (here x is 1,2,3 or 4) from all your .TXT files located in the folder _TXT-TREE (sub-directories allowed) and compares (phrase-checks) them against respective Gamera x-gram corpus creating four files: GALFONx_overlapped_all_digits_distinct.txt; GALFONx_overlapped_all_distinct.txt; GALFONx_progenitor_all_distinct.txt; GALFONx_unfamiliar_all_distinct.txt

When executing e.g. 'GAMERA P.C. 3-grams via sort' icon **GALFON3_overlapped_all_digits_distinct.txt** is auto-loaded into NOTEPAD:

```
...
enjoys_widespread_adoption      0,000,002      enough_bandwidth_to      0,000,112
enormous_bigger_than            0,000,001      will_occasionally_forget  0,000,005
...
```

Number and size of all Gamera .TXT files (master-corpus): D:\>dir _Gamera*.txt /og/on/s/b/a-d gives **562,504** files in **33,750,715,417** bytes, 6,528 folders.