

# Business information consumption: 9,570,000,000,000,000,000,000 bytes per year

April 6 2011

Byte (B)	=	1 byte	=	1	=	One character of text
Kilobyte (KB)	=	10 <sup>3</sup> bytes	=	1,000	=	One page of text
Megabyte (MB)	=	10 <sup>6</sup> bytes	=	1,000,000	=	One small photo
Gigabyte (GB)	=	10 <sup>9</sup> bytes	=	1,000,000,000	=	One hour of High-Definition video, recorded on a digital video camera at its highest quality setting, is approximately 7 Gigabytes
Terabyte (TB)	=	10 <sup>12</sup> bytes	=	1,000,000,000,000	=	The largest consumer hard drive in 2008
Petabyte (PB)	=	10 <sup>15</sup> bytes	=	1,000,000,000,000,000	=	AT&T carried about 18.7 Petabytes of data traffic on an average business day in 2008
Exabyte (EB)	=	10 <sup>18</sup> bytes	=	1,000,000,000,000,000,000	=	Approximately all of the hard drives in home computers in Minnesota, which has a population of 5.1M
Zettabyte (ZB)	=	10 <sup>21</sup> bytes	=	1,000,000,000,000,000,000,000	=	

(PhysOrg.com) -- Three scientists at UC San Diego have rigorously estimated the annual amount of business-related information processed by the world's computer servers in terms that Gutenberg and Galileo would have appreciated: the digital equivalent of a 5.6-billion-mile-high stack of books from Earth to Neptune and back to Earth, repeated about 20 times a year.

The world's roughly 27 million [computer servers](#) processed 9.57 zettabytes of information in 2008, according to a paper to be presented April 7 at Storage Networking World's (SNW's) annual meeting in Santa

Clara, Calif.

The first-of-its kind rigorous estimate was generated with server-processing performance standards, server-industry reports, interviews with information technology experts, sales figures from server manufacturers and other sources. (One zettabyte is 10 to the 21st power, or a million million gigabytes.)

The study estimated that enterprise server workloads are doubling about every two years, which means that by 2024 the world's enterprise servers will annually process the digital equivalent of a stack of books extending more than 4.37 light-years to Alpha Centauri, our closest neighboring star system in the [Milky Way](#) Galaxy. (Each book is assumed to be 4.8 centimeters thick and contain 2.5 megabytes of information.)

"Most of this information is incredibly transient: it is created, used, and discarded in a few seconds without ever being seen by a person," said Roger Bohn, one of the report's co-authors and a professor of technology management at UC San Diego's School of International Relations and Pacific Studies. "It's the underwater base of the iceberg that runs the world that we see."

The authors of the report titled "How Much Information: 2010 Report on Enterprise Server Information" are Bohn, James E. Short, a research scientist at UC San Diego's School of International Relations and Pacific Studies and research director of the HMI? project, and Chaitanya K. Baru, a distinguished scientist at the San Diego Supercomputer Center.

**Table 1: World Server Information**

Summary of Results		
What is measured	World 2008 Total	Notes
Bytes processed plus delivered	9.57 zettabytes	
Bytes per worker per year	3.01 terabytes	3.18 billion workers in world labor force
Bytes per worker per day	12.0 gigabytes	
Bytes per company per year	63.4 terabytes	151 million world businesses registered

Sources: CIA Factbook, World Labor Organization, Dun & Bradstreet

The paper follows an earlier report on information consumption by U.S. households as part of The How Much Information? project. The effort is designed to conduct a census of the world's information in 2008 and onward, and is supported by AT&T, Cisco Systems, IBM, Intel, LSI, Oracle and Seagate Technology. Early support was provided by the Alfred P. Sloan Foundation.

"The exploding growth in stored collections of numbers, images and other data is well known, but mere data becomes more important when it is actively processed by servers as representing meaningful information delivered for an ever-increasing number of uses," said Short. "As the capacity of servers to process the digital universe's expanding base of information continues to increase, the development itself creates

unprecedented challenges and opportunities for corporate information officers."

The workload of all 27 million of the world's enterprise servers in use in 2008 was estimated by using cost and performance benchmarks for online transaction processing, Web services and virtual machine processing.

"Of course, we couldn't directly measure the allocation of workload to millions of servers worldwide, but we received important guidance from experts, industry data and our own judgment," Short said. "Since our capacity assumptions, methodology and calculations are complex, we have prepared a separate technical paper as background to explain our methodology and provide sample calculations."

Servers amount to the unseen, ubiquitous, humming computational infrastructure of modern economies. The study estimated that each of the 3.18 billion workers in the world's labor force received an average of 3 terabytes of information per year.

Rather than focusing on raw processing power, the new analysis focused on server performance per dollar invested as a more consistent yardstick across a wide array of server types and sizes. "While midrange servers doubled their Web processing and business application workloads every 2 years, they doubled their performance per dollar every 1.5 years," Bohn said.

The 36-page "How Much Information?" report said total worldwide sales of all servers has remained stable at about \$50-\$55 billion per year for five years ending in 2008, while new-server performance as measured by industry benchmarks went up five- to eight-fold during the same period. Entry-level servers costing less than \$25,000 processed about 65 percent of the world's information in 2008, midrange servers processed 30

percent, and high-end servers costing \$500,000 or more processed 5 percent of the world's information in 2008.

The report's authors note that the estimated workload of the world's servers may be an underestimate because server-industry sales figures don't fully include the millions of servers built in-house from component parts by Google, Microsoft, Yahoo! and others.

The study estimated a sharp increase in virtualization beginning in 2006, in which many distinct "virtual servers" can run on one physical server. Virtualization is a way to improve energy efficiency, scalability and overall performance of large-scale information processing. One of its uses is for cloud computing in which server-processing power is provided as a centrally administered commodity that business clients can pay for as needed.

"Corporations and organizations that have huge and growing databases are compelled to rethink how they accomplish economies of scale, which is why many are now embracing cloud computing initiatives and green datacenters," said Baru. "In addition, a corporation's competitiveness will increasingly hinge on its ability to employ innovative search techniques that help users discover data and obtain useful results, and automatically offer recommendations for subsequent searches."

Measuring worldwide flows of information is an inexact science, and the How Much Information project will issue additional analyses as improved metrics become available and accepted. In 2007, the International Data Corporation and EMC Corp. reported that the total digital universe of information created, captured or replicated digitally was 281 exabytes and would not reach 1 [zettabyte](#) until 2010. The study by Short, Bohn and Baru included estimates of the amount of data processed as input and delivered by servers as output. For example, one email message may flow through multiple servers and be counted

multiple times.

The How Much [Information](#) paper points to the importance of data archiving and digital-data preservation. "Preserving data is an increasingly important challenge for business organizations and arbitrary age limits make little sense," said Baru. "In the future, data archiving and preservation will require as much enthusiasm in research and industry settings as we have provided to data generation and data processing."

Provided by University of California - San Diego

Citation: Business information consumption: 9,570,000,000,000,000,000 bytes per year (2011, April 6) retrieved 26 April 2024 from <https://phys.org/news/2011-04-world-consumption-bytes-year.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--