The Diverse and Exploding Digital Universe

Executive Summary

An Updated Forecast of Worldwide Information Growth Through 2011

March 2008

John F. Gantz, Project Director

Christopher Chute
Alex Manfrediz
Stephen Minton
David Reinsel
Wolfgang Schlichting
Anna Toncheva

IDC
Analyze the Future

EXECUTIVE SUMMARY

This white paper, sponsored by EMC, is an update of IDC's inaugural forecast of the digital universe published in March 2007. In this year's update we calibrate the size (bigger) and growth (faster) of the digital universe again, but we also explore some areas we only touched on last time. As before, we also seek to understand the implications for business, government, and society.

Some key findings are as follows:

- The digital universe in 2007 — at 2.25×10^{21} bits (281 exabytes or 281 billion gigabytes) — was 10% bigger than we thought. The resizing comes as a result of faster growth in cameras, digital TV shipments, and better understanding of information replication.

- By 2011, the digital universe will be 10 times the size it was in 2006.

- As forecast, the amount of information created, captured, or replicated exceeded available storage for the first time in 2007. Not all information created and transmitted gets stored, but by 2011, almost half of the digital universe will not have a permanent home.

- Fast-growing corners of the digital universe include those related to digital TV, surveillance cameras, Internet access in emerging countries, sensor-based applications, datacenters supporting "cloud computing," and social networks.

- The diversity of the digital universe can be seen in the variability of file sizes, from 6 gigabyte movies on DVD to 128-bit signals from RFID tags. Because of the growth of VoIP, sensors, and RFID, the number of electronic information "containers" — files, images, packets, tag contents — is growing 50% faster than the number of gigabytes. The information created in 2011 will be contained in more than 20 quadrillion — 20 million billion — of such containers, a tremendous management challenge for both businesses and consumers.

- Of that portion of the digital universe created by individuals, less than half can be accounted for by user activities — pictures taken, phone calls made, emails sent — while the rest constitutes a digital "shadow" — surveillance photos, Web search histories, financial transaction journals, mailing lists, and so on.

- The enterprise share of the digital universe is widely skewed by industry, having little relationship to GDP or IT spending. The finance industry, for instance, accounts for almost 20% of worldwide IT spending but only 6% of the digital universe. Meanwhile, media, entertainment, and communications industries will account for 10 times their share of the digital universe in 2011 as their share of worldwide gross economic output.

- The picture related to the source and governance of digital information remains intact: Approximately 70% of the digital universe is created by individuals, but enterprises are responsible for the security, privacy, reliability, and compliance of 85%.

To deal with this explosion of the digital universe in size and complexity, IT organizations will face three main imperatives:

One. They will need to transform their existing relationships with the business units. It will take all competent hands in an organization to deal with information creation, storage, management, security, retention, and disposal in an enterprise. Dealing with the digital universe is not a technical problem alone.

Two. They will need to spearhead the development of organizationwide policies for information governance: information security, information retention, data access, and compliance.

Three. They will need to rush new tools and standards into the organization, from storage optimization, unstructured data search, and database analytics to resource pooling (virtualization) and management and security tools. All will be required to make the information infrastructure as flexible, adaptable, and scalable as possible.

We have many of the tools in place — from Web 2.0 technologies and terabyte drives to unstructured data search software and the Semantic Web — to tame the digital universe. Done right, we can turn information growth into economic growth.

METHODOLOGY

Our basic approach of sizing the digital universe was to:

- Develop a forecast for the installed base of any of 30 or so classes of device or application that could capture or create digital information.

- Estimate how many units of information — files, images, songs, minutes of video, calls per capita, packets of information — were created in a year.

- Convert the units of information to megabytes using assumptions about resolutions, compression, and usage.

- Estimate the number of times a unit of information might be replicated, either to share or store.

Much of this information is part of IDC's ongoing research (see the Bibliography). Figure 1 provides a list of the kinds of devices or information categories we examined.

Figure 1

Devices and Applications Tracked

Source: IDC, 2008

IDC

Analyze the Future

Global Headquarters:
5 Speen Street • Framingham, MA 01701
508.872.8200
www.idc.com