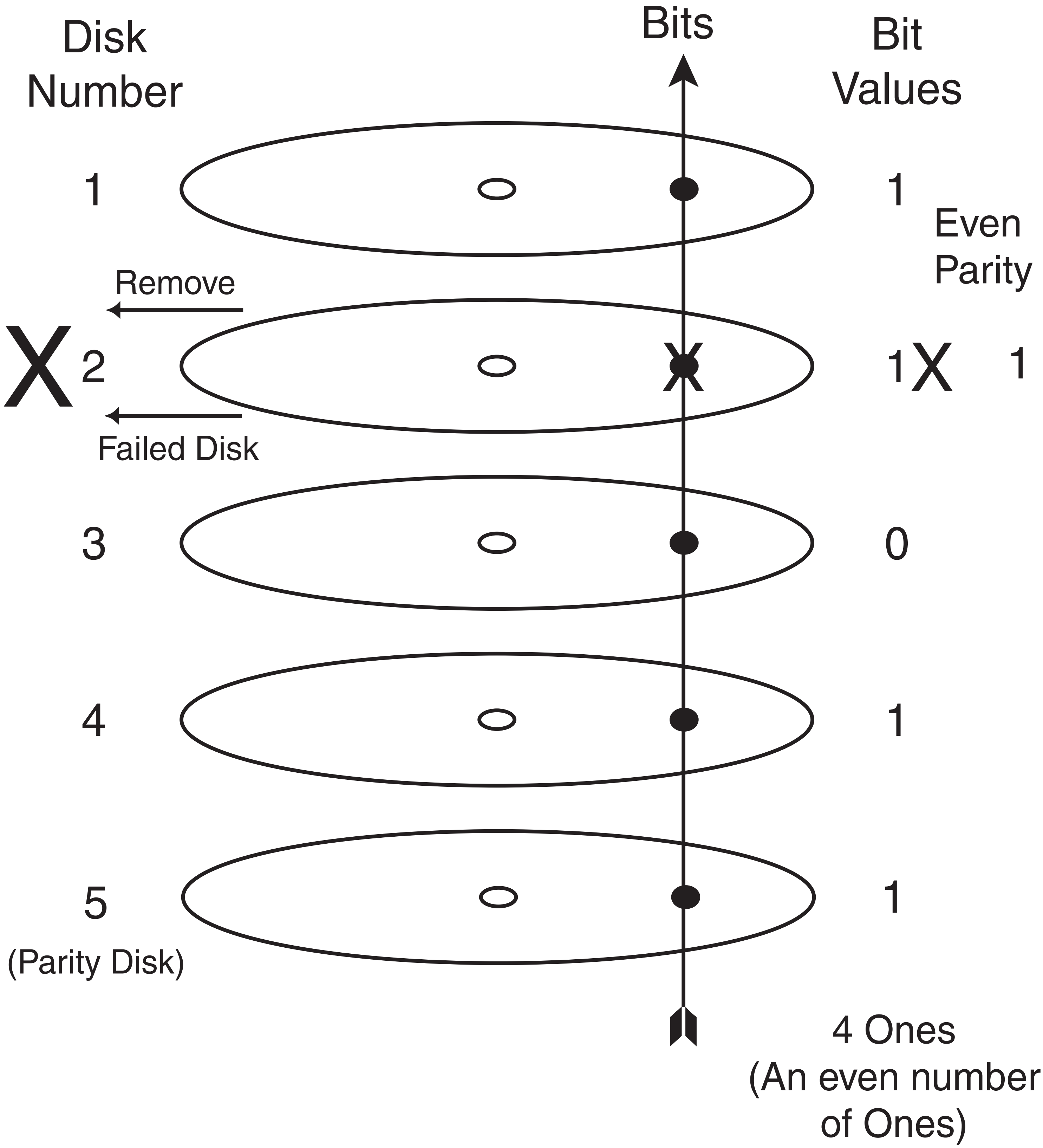
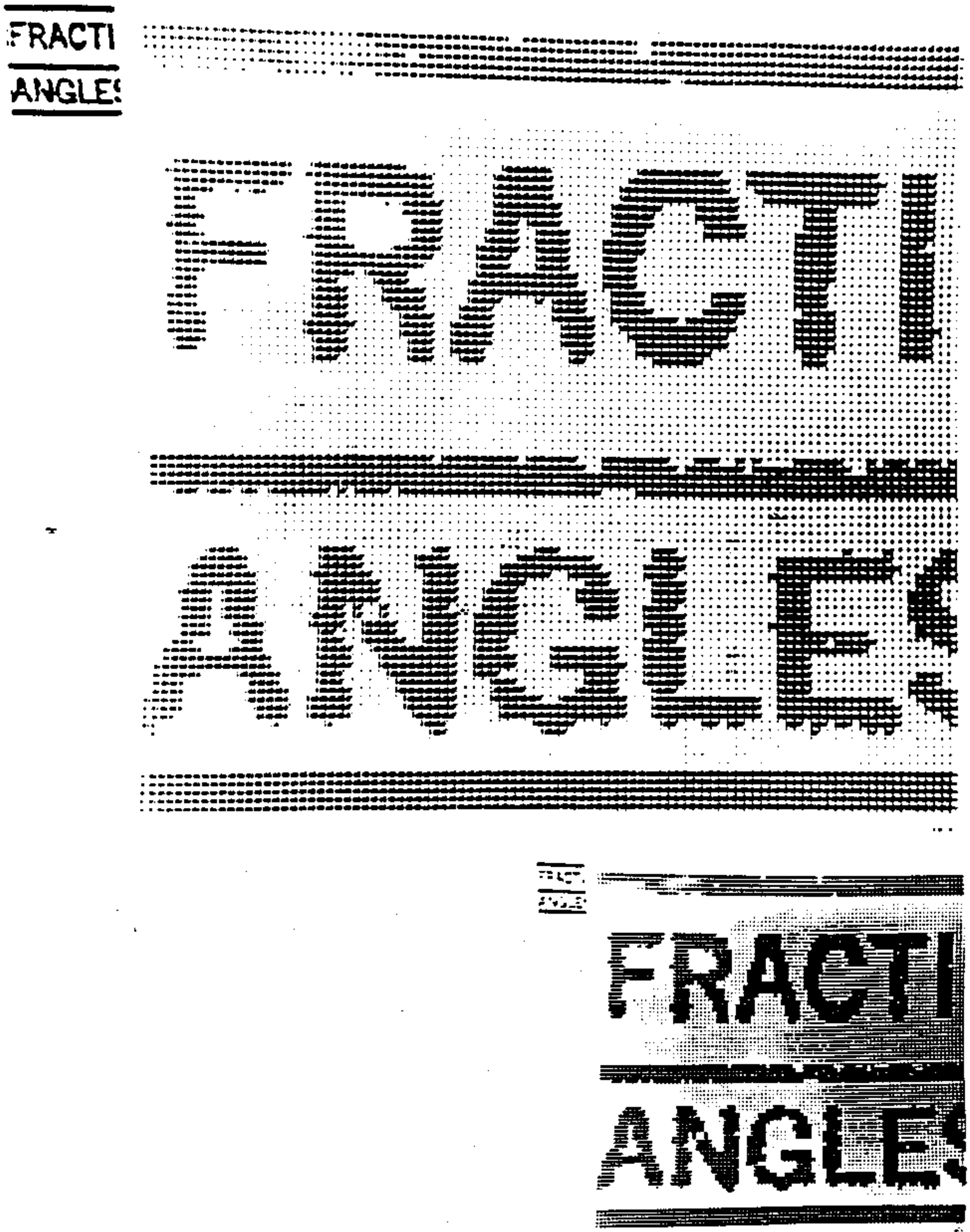
RAID (Redundant Array of Inexpensive magnetic Disks)

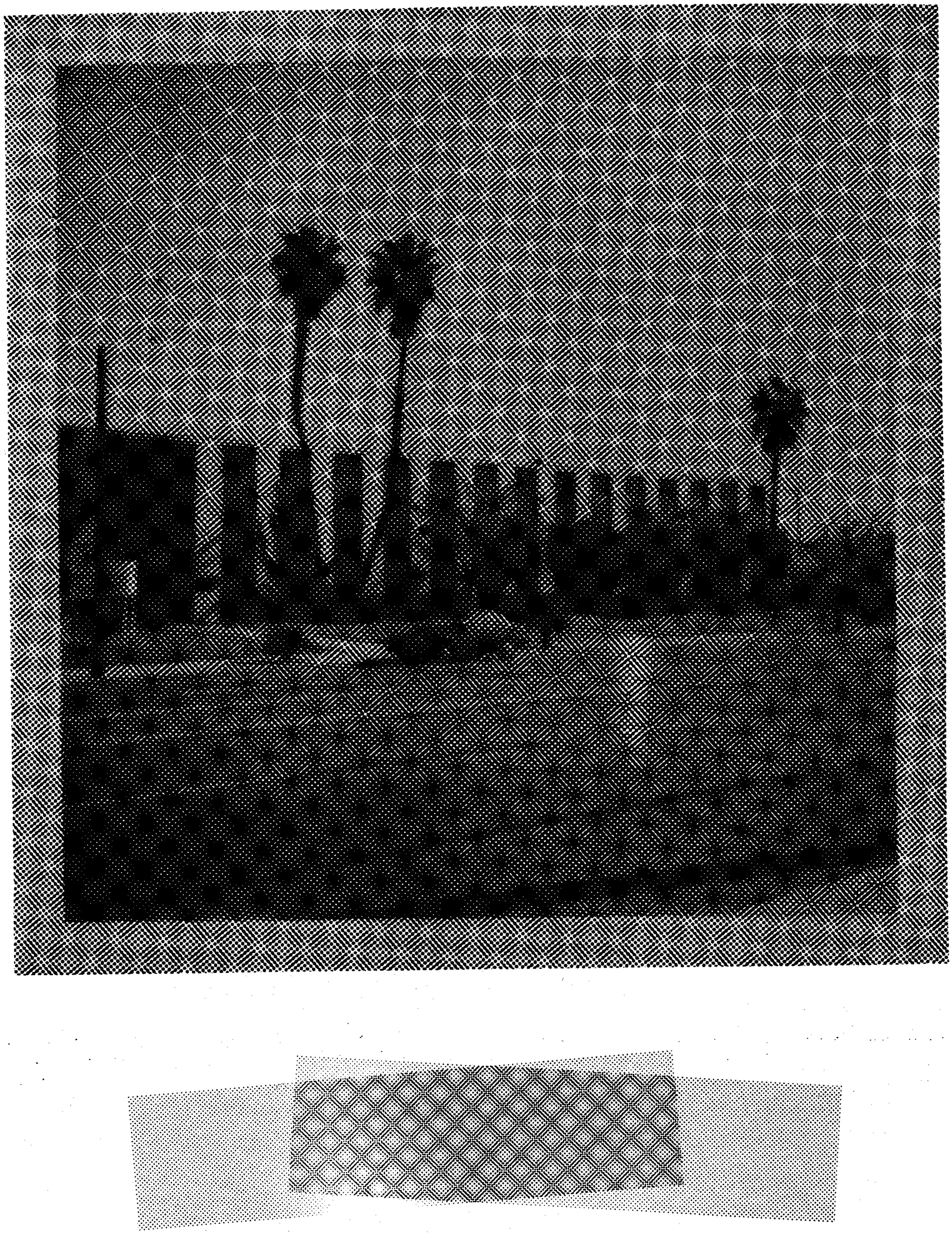


An array of 5 identical magnetic disks



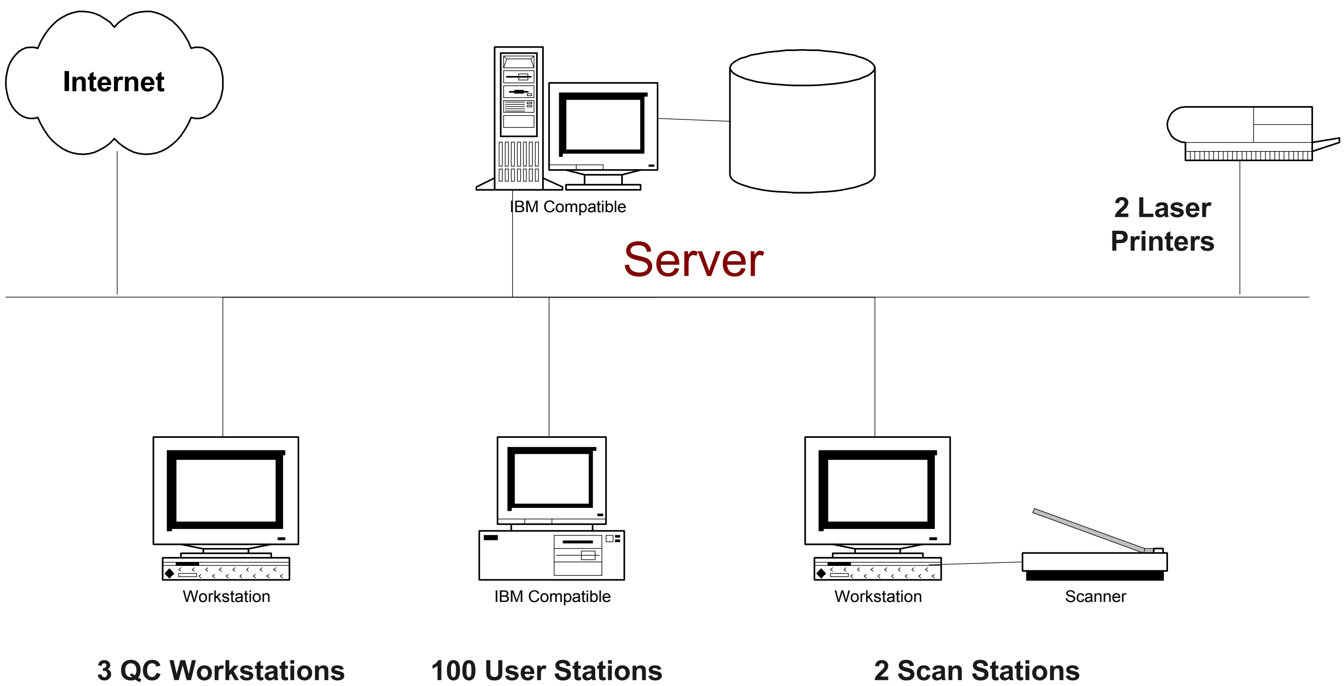
Y-1 Raster Scanned Characters

Y-10 Halftoned Image



Y-11 Rescreened Halftone Image (Moiré Pattern)

Network Diagram



Y-17 **Network Diagram (Portrait)**

```
==================== Start of EPS (Encapsulated PostScript) ================
=================== PDL (Page Description Language) Program ================

%%Creator: Adobe Illustrator(R) 8.0
%%AI8_CreatorVersion: 8
%%For: (Steve Gilheany) ( )
%%Title: (22021v026  How Digitizing Works for Non-Technical Manager.eps)
%%CreationDate: (3/19/00) (11:13 AM)
%%BoundingBox: -222 -840 1088 775
%%HiResBoundingBox: -221.4863 -839.75 1087.6953 774.5938
%%DocumentProcessColors: Black
%%DocumentFonts: Helvetica

========================= [ 3 pages of text elided* ]  ======================

            } bind def
            mark
            /setcustomcolor where not
            {
                        /findcmykcustomcolor
                        {
                                    (AI8_CMYK_CustomColor)
                                    6 packedarray
                        } bind def
                        /findrgbcustomcolor
                        {
                                    (AI8_RGB_CustomColor)
                                    5 packedarray
                        } bind def
                        /setcustomcolor
                        {
                                    exch
                                    aload pop dup
                                    (AI8_CMYK_CustomColor) eq
                                    {
                                                pop pop
                                                4
                                                {
                                                            4 index mul
                                                            4 1 roll
                                                } repeat
                                                5 -1 roll pop
                                                setcmykcolor
                                    }
                                    {
                                                dup (AI8_RGB_CustomColor) eq
                                                {
                                                            pop pop
                                                            3
                                                            {
                                                                  1 exch sub
                                                                  3 index mul
                                                                  1 exch sub
                                                                  3 1 roll
                                                            } repeat
                                                            4 -1 roll pop
                                                            setrgbcolor
                                                }
                                                {
                                                            pop
                                                            4
                                                            {
                                                                  4 index mul 4 1
                                                                    [fixed] roll
                                                            } repeat
                                                            5 -1 roll pop
                                                            setcmykcolor
                                                } ifelse
                                    } ifelse
                        }
                        def
            } if
            /setAIseparationgray
```

```
=========== [ 81 pages of text elided* ]  ===========

1 0 0 1 398.8359 708.1401 0 Tp
0 Tv
TP
0 Tr
0 0 0 1 k
/_Helvetica 18.3526 17.0863 -4.1292 Tf
114.0882 100 Tz
(Photon) Tx 1 0 Tk     <<=========== Photon ==========
(\r) TX
TO
0 To
1 0 0 1 485.0859 591.1401 0 Tp
0 Tv
TP
0 Tr
(Electron) Tx 1 0 Tk    <<========== Electron ========
(\r) TX
TO
1 Ap
464.1475 597.75 m
464.1475 594.3247 460.9805 591.5493 457.0723 591.5493 c
453.167 591.5493 449.999 594.3247 449.999 597.75 c
449.999 601.1733 453.167 603.9497 457.0723 603.9497 c
460.9805 603.9497 464.1475 601.1733 464.1475 597.75 c
f
436.6592 695.5898 m
436.6592 692.1655 433.4932 689.3896 429.5859 689.3896 c
425.6797 689.3896 422.5127 692.1655 422.5127 695.5898 c
422.5127 699.0142 425.6797 701.7905 429.5859 701.7905 c
433.4932 701.7905 436.6592 699.0142 436.6592 695.5898 c
f
0 To
1 0 0 1 92.1323 541.9849 0 Tp
0 Tv
TP
0 Tr
(P) Tx 1 40 Tk    <<============ Paper =============
(aper) Tx 1 0 Tk
(\r) TX
TO
0 To
1 0 0 1 267.2729 646.1987 0 Tp
0 Tv
TP
0 Tr
(Lens) Tx 1 0 Tk    <<============= Lens ===========
(\r) TX
TO
0 To
1 0 0 1 146.5229 698.6987 0 Tp
0 Tv
TP
0 Tr
/_Helvetica 19.2467 17.9187 -4.3303 Tf
108.7885 100 Tz
(CCD) Tx 1 0 Tk     <<============= CCD ============
(\r) TX
TO
0 To
1 0 0 1 61.4785 618.4546 0 Tp
0 Tv

=============== [ 6 pages of text elided ] ====================

Adobe_typography_AI5 /terminate get exec
Adobe_cshow /terminate get exec
Adobe_level2_AI5 /terminate get exec
%%EOF

============== End of EPS  PDL Program =============

[*Count of pages elided is for lettersize, single column in this font and size.]
```

Paper 22025 Figure 13. A portion of the document description (written in PostScript) for the image from *How Digitizing Works* in Figure 12.

Bibliography

(Paper 22025) (Most titles followed by URLs (Universal Resource Locators) are available free on the Internet)

Adobe Systems, *PostScript Language Reference*, Third Edition, Addison-Wesley, Menlo Park, CA, 1999, ISBN 0-201-37922-8 [http://www.adobe.com/products/postscript/pdfs/PLRM.pdf]

Adobe Systems, *PostScript Language Reference Supplement,* Adobe PostScript 3, Version 3010 and 3011, Product Supplement, 1999, 28-007 [http://partners.adobe.com/asn/developer/PDFS/TN/PS3010and3011.Supplement.pdf]

Adobe Systems, Inc., *Adobe Type 1 Font Format,* Addison-Wesley, Menlo Park, CA, 1990 ISBN 0-201-57044-0 [http://partners.adobe.com/asn/developer/PDFS/TN/T1_SPEC.PDF]

Adobe Systems, *Type 1 Font Format, Supplement,* 1994, Technical Specification #5015, PN (Part Number) LPS5015 [http://partners.adobe.com/asn/developer/PDFS/TN/5015.Type1_Supp.pdf]

Adobe Systems, *Portable Document Format Reference Manual*, Version 1.3, Addison-Wesley, Menlo Park, CA, 1999, ISBN 0-201-62628-4 [http://partners.adobe.com/asn/developer/acrosdk/DOCS/pdfspec.pdf]

Adobe Systems, *Display Postscript Manual*, 1993 [http://partners.adobe.com/asn/developer/PDFS/TN/DPS.refmanuals.PSW.pdf]

Adobe Extreme, brochure (integrates PDF and Postscript Printing), 1998, Adobe Systems [http://www.adobe.com/products/extreme/pdfs/extremewp.pdf]
Ried, Glenn C., *PostScript Language Program Design*, Addison-Wesley, Menlo Park, CA, ISBN 0-201-14396-8 1988
Adobe Systems, *PostScript Language Tutorial and Cookbook* Addison-Wesley Menlo Park, CA, 1986 ISBN 0-201-10189-0
Adobe Systems, *PostScript Language Document Structuring Conventions Specification*, Version 3.0, Adobe Systems, 1992, PN LPS5001 [http://partners.adobe.com/asn/developer/PDFS/TN/5001.DSC_Spec.pdf]

Microfilm life expectancy of 500 years - *ANSI/NAPM IT9.1-1992 Imaging Media (Film)-Silver-Gelatin Type-Specifications for Stability* gives the maximum concentration of residual thiosulfate in microfilm that will allow for a microfilm life expectancy of 500 years. (American National Standards Institute [http://www.ANSI.org], National Association of Photographic Manufacturers [http://www.techexpo.com/tech_soc/napm.html]). For film storage requirements, see Kodak.com at [http://www.kodak.com/cluster/global/en/consumer/products/techInfo/e30/e30Contents.shtml]

Jeff Rothenberg, 'Ensuring the Longevity of Digital Documents', *Scientific American*, January 1995, Vol 272, No 1, pages 42-47

Harry Nyquist, "Certain Topics in Telegraph Transmission Theory," *Trans.*, AIEEE, Vol. 47, April 1928, pp. 617-644.

George Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", *The Psychological Review,* 1956, 63:81-97. The text is available on the Internet at: [http://cogprints.soton.ac.uk/documents/disk0/00/00/07/30/cog00000730-00/miller.html]

ISO 9000 International Quality Standard (International Standards Organization) [http://www.iso.ch/9000e/9k14ke.htm]

Matrix, Warner Brothers, 1999; *Blade Runner*, Warner Brothers, 1982 [http://www.WB.com]

The *Rosetta Stone* is part of the collection of the British Museum, London, collection number EA 24. "[The Rosetta Stone] ends by saying that it is to be made known (in March, 196 BC, Before Christ) that all the men {people} of Egypt should magnify and honor [King] Ptolemy V, and that the text should be set up in hard stone in the three scripts which the Rosetta Stone still bears today (hieroglyphic, Demotic, and Greek)". Thus, the Rosetta Stone contains its own metadata, and a single document reproduced in three formats that are locked (sealed) together in stone, for the purpose of causing the message to last a long period of time; like a permanent virtual fascicle. The following provides a Rosetta Stone history: [http://www.thebritishmuseum.ac.uk/egyptian/ea/gall/rosetta.html] The following provides an explanation of the Rosetta Stone text: [http://www.thebritishmuseum.ac.uk/egyptian/ea/further/rosettasay.html]

Customary Units	**Number**	**Equivalent Customary Units**	**Number of Common Units**	**Common Units**	**Number**	**Hertz Range**	**Power of 10**
electron frequency (at C, the speed of light)	2.454	picometers	1,200,000,000,000,000,000	hertz	1.20	exahertz	18
fiber optic wavelength (carrier frequency)	1,500	nanometers (=1.5 um)	230,000,000,000,000	hertz	230.00	terahertz	12
microprocessor clock rate (cycle time)	1	billion clock cycles/sec.	1,000,000,000	hertz	1.00	gigahertz	9
Computer RAM (Random Access Memory)	50	nanoseconds	20,000,000	hertz	20.00	megahertz	6
magnetic disk access time (12,000 RPM)	5	milliseconds	200	hertz	200.00	hertz	0
jukebox access (picker) (1 to 5 seconds)	1	second	1	hertz	1.00	hertz	0
second (1 hertz = 1 cycle per second)	1	second	1	seconds	1.00	hertz	0
minute	60	seconds	60	seconds	16.67	millihertz	-3
hour	60	minutes	3,600	seconds	277.78	microhertz	-6
day	24	hours	86,400	seconds	11.57	microhertz	-6
week	7	days	604,800	seconds	1.65	microhertz	-6
month	1/12	year	2,629,800	seconds	380.26	nanohertz	-9
year	365.25	days	31,557,600	seconds	31.69	nanohertz	-9
year	1	year	1	years	31.69	nanohertz	-9
decade	1	ten years	10	years	3.17	nanohertz	-9
century	1	hundred years	100	years	316.90	picohertz	-12
millennium	1	million years	1,000	years	31.69	picohertz	-12
precession of the Zodiac	1	rotation	26,000	years	1.22	picohertz	-12
million years	1	million years	1,000,000	years	31.69	femtohertz	-15
billion years	1	billion years	1,000,000,000	years	31.69	attohertz	-18
period of the universe (postulated)	1	period	85,000,000,000	years	372.80	zeptohertz	-21

Paper 22025 Figure 21. Inverse Table of Periodicity

Projecting the Cost of Magnetic Disk Storage Over the Next 10 Years

The following table shows the cost per GigaByte and the cost per TeraByte (one thousand GigaBytes) in each year from 1992 through the year 2010. As shown in the table, one GigaByte is sufficient storage for two file cabinets of scanned documents and one TeraByte is sufficient storage for two thousand file cabinets of scanned documents. The table is generated using an estimated price reduction assumption of 45 percent each year. Disk configurations, of varying complexity, are given for each year to explain possible variations in quoted prices.

Annual Decline	Cost For 1 GigaByte	Cost For 1 TeraByte = 1,000 GigaBytes (US Dollars) (Storage for 2,000 Scanned File Cabinets) (Holding 20 Million Scanned Letter Size Pages)				
45%	1,000 MBytes	Non-FC/SCSI	Non-FC/SCSI	SAN	SCSI/FC SAN/PC	
Year	(US Dollars) (Storage for 2 Scanned File Cabinets)	PC Disk No Online Redundancy	PC Disk Software RAID Redundancy	FC Disk FC Fabric Hardware RAID	Name Brand Fault Awareness Hardware RAID	Mainframe
		1 X	2 X	4 X	8 X	12 X
1992	1,000.00	1,000,000.00	2,000,000.00	4,000,000.00	8,000,000.00	12,000,000.00
1993	550.00	550,000.00	1,100,000.00	2,200,000.00	4,400,000.00	6,600,000.00
1994	302.50	302,500.00	605,000.00	1,210,000.00	2,420,000.00	3,630,000.00
1995	166.38	166,375.00	332,750.00	665,500.00	1,331,000.00	1,996,500.00
1996	91.51	91,506.25	183,012.50	366,025.00	732,050.00	1,098,075.00
1997	50.33	50,328.44	100,656.88	201,313.75	402,627.50	603,941.25
1998	27.68	27,680.64	55,361.28	110,722.56	221,445.13	332,167.69
1999	15.22	15,224.35	30,448.70	60,897.41	121,794.82	182,692.23
2000	8.37	8,373.39	16,746.79	33,493.58	66,987.15	100,480.73
2001	4.61	4,605.37	9,210.73	18,421.47	36,842.93	55,264.40
2002	2.53	2,532.95	5,065.90	10,131.81	20,263.61	30,395.42
2003	1.39	1,393.12	2,786.25	5,572.49	11,144.99	16,717.48
2004	0.77	766.22	1,532.44	3,064.87	6,129.74	9,194.61
2005	0.42	421.42	842.84	1,685.68	3,371.36	5,057.04
2006	0.23	231.78	463.56	927.12	1,854.25	2,781.37
2007	0.13	127.48	254.96	509.92	1,019.84	1,529.75
2008	0.07	70.11	140.23	280.45	560.91	841.36
2009	0.04	38.56	77.13	154.25	308.50	462.75
2010	0.02	21.21	42.42	84.84	169.68	254.51

The Basis for the Assumptions in the Table

The above decreases in price are based on estimated increases in disk storage density by IBM, which are based on IBM's predictions for its magnetoresistive (MR) head technology. IBM invented the MR technology, and MR is currently the technological basis for advancements in the magnetic disk industry. IBM had been increasing the areal bit density of magnetic disks at a rate of 60 percent per year from 1989 to 1994. In 1994, IBM projected that the 60 percent rate of increase would continue for the foreseeable future. (Source: The Era of Magnetoresistive Heads, Ed Grochowski, IBM Research Division, Almaden Research Center, San Jose, CA., 1994).

With IBM's projected rate of increase in areal bit density, of 60 percent per year, for a given price and a given year, one could purchase 1.6 times as much storage capacity the following year. This corresponded to a constant decrease in the price of magnetic storage of 37.5 percent per year.

In a press release issued on December 29, 1997, IBM stated that the percent of price decrease was continuing on track. On October 4, 1999 IBM issued a press release stating that the rate of increase in disk storage density had increased from 60 to 100 percent per year in each of the last two years. IBM also announced that it had demonstrated very stable bit densities of 35.3 billion data bits per square inch in the lab. This could soon lead to the manufacture of 500 GigaByte 3 1/2 inch form factor magnetic disk drives (2 of the 500 GigaByte drives would store 1 TeraByte). IBM anticipated that the increase in density would continue. On October 15, 1999, IBM announced a 73 GigaByte, 3.5 inch, multi-platter, disk with a 2 Gigabit per second, serial, fiber channel, interface.

Since 1992, when 1 GigaByte cost 1 thousand US dollars, and 1 TeraByte was too expensive for most applications, at 1 million US dollars, memory costs have declined at about 45 percent per year. This is the percent decline used to project the next ten years in the chart above. A decline of 45 percent per year is slower than the 50 percent per year decline for the last two years (1998-2000), and slightly faster than the average decline of more than 40 percent over the preceding 6 years (From 1992 to 1998).

A Longer Perspective for Retention Periods

IBM introduced the 5 MegaByte RAMAC disk drive in June 1957, at a monthly rental of US$ 3,200.00 (in 1957 dollars). (Source: IBM's Early Computers, by Charles J. Bashe, MIT Press, Cambridge, MA, 1986.) IBM shipped the first RAMAC (Random Access Method of Accounting and Control) System magnetic disks (known historically as RAMAC disks) to Zellerbach Paper in San Francisco. The RAMAC disks cost 100 thousand US dollars per MegaByte, or 100 million US dollars per GigaByte, to purchase. (Adjusted to current US dollars, and adjusted for IBM's historic practice of renting rather than selling.) In 2000, magnetic disk storage cost 8 US dollars per GigaByte. This represents a decline of over ten-million-to-one (from 100 million US dollars to less than 10 US dollars per MegaByte) in forty-three years, or a price decline of about 31.6 percent e per year.

The advances described here are based on magnetic disk technology. The study of the history of technology has shown that over long periods of time, as older technologies are exhausted, new technologies replace them, and a steady rate of advancement is maintained. Holographic and molecular machine nano-technology (see [http://www.Foresight.org]) have been under development for some time and promise several more orders of magnitude improvement in price and physical size reduction, as well as increases in speed, when advances in magnetic disk technology slow.

Cost Adjustment for Advanced Hardware and Support Technology

There are many configurations of magnetic disks available. The simplest disk configuration is in the PC (Personal Computer) on one's desk. In the preceding table, this configuration is assigned an approximate relative cost of 1X. 1X is the base cost for the cost comparison of the disk configurations. The next configuration adds redundancy by storing the same data on two or more disks. This is assigned a relative approximate cost of 2X because the disk storage cost is roughly twice as much as the disk storage cost in a generic PC. The third configuration adds hardware support and fault awareness to the redundancy in a SAN (Storage / System Area Network) configuration using FC (Fiber Channel) interfaces. This is assigned a relative approximate cost of 4X. The fourth configuration adds hardware support, fault awareness, and name branding to the redundancy using either a name brand high end server PC or a name brand SAN using SCSI (Small Computer System Interface) or FC interfaces. This is assigned a relative approximate cost of 8X. Finally, mainframe disk configurations add more hardware and software features to data storage, again with name branding, resulting in a relative approximate cost of 12X. (All trademarks are the property of their respective holders.)

Compiled by Steve Gilheany, CRM, CDIA

Digital Image Sizes

Computer storage requirements for various digitized document types (Estimates are rounded and adjusted for ease of use.) **(Page 1 of 6)**

Scanned Letter Size Pages

(All images are scanned 1 bit per pixel, black & white, and compressed, unless otherwise noted.)

1 scanned page (8 1/2 by 11 inches, A4) = 50 KiloBytes (KByte) (on average, black & white, CCITT G4 compressed)
1 file cabinet (4 drawer) (10,000 pages on average) = 500 MegaBytes (MByte) = 1 CD (Compact Disc) (ROM or WORM)
2 file cabinets = 10 cubic feet (cf) = 1,000 MegaBytes = 1 GigaByte (GByte); 10 file cabinets = 1 DVD-R (WORM) (see below)
2,000 file cabinets = 1,000 GigaBytes = 1 TeraByte (TByte); 2,000 file cabinets = 200 DVDs
1 box (in inches: 15 1/2 long x 12 wide x 10 deep) (400 x 300 x 250 mm) (2,500 pages) = 1 file drawer = 125 MegaBytes
1 box (packed) = 2 linear feet (500 mm) of files (loose enough for active filing) = 25 (rounded) linear inches = 125 MegaBytes
1 linear inch (~20 mm) = 100 pages = 5 MegaBytes; 1 thousand linear inches = 100 thousand pages = 5 GigaBytes
1 cubic foot (cf) (~.025 cubic meter) = 2000 pages = 100 MegaBytes; 10 cubic feet (~.25 cubic meters) = 20 thousand pages = 1 GigaBytes
8 boxes = 16 linear feet = 2 file cabinets = 1 GigaByte; 8,000 boxes = 16,000 linear feet = 1,000 GigaBytes = 1 TeraByte

For paper and microform document imaging, see also AIIM (Association for Information and Image Management) [http://www.AIIM.org]
For records and information management, see also ARMA (Association of Records Managers and Administrators) [http://www.ARMA.org]

Scanned Engineering Drawings / Large Format Documents

1 E size drawing (48 inches by 36 inches) (A0 size) = 16 letter size pages (8 1/2 by 11 inches, metric A4) = 800 KiloBytes To place an E size drawing in a file folder in a file cabinet drawer, the drawing must be folded in half 4 times and is then 16 sheets of paper thick when folded.

NB: Scanning must accommodate the older, untrimmed, US paper sizes, because it is the older drawings that are digitized by scanning.

Metric Trimmed Paper Sizes					United States Paper Sizes				Storage
Metric Name	**Metric Size in Millimeters**	**Size in Inches**	**Number of Square Meters**	**Number of A4 Size Pages**	**US Name**	**New Size (Trimmed) in Inches**	**Old Size (Untrimmed) in Inches**	**Equivalent Letter Size Pages**	**Digital Image Storage Requirements**
A8	52 x 74	2.07 x 2.91	1 / 256	1 / 16	Business Card				5 KiloBytes
A7	74 x 105	2.91 x 4.13	1 / 128	1 / 8	3 x 5	3 x 5			10 KiloBytes
A6	105 x 148	4.13 x 5.83	1 / 64	1 / 4	Microfiche				
A5	148 x 210	5.83 x 8.27	1 / 32	1 / 2	5 x 8	5 x 8			25 KiloBytes
A4	210 x 297	8.27 x 11.69	1 / 16	1	A	8 1 / 2 x 11	9 x 12	1	50 KiloBytes
A3	297 x 420	11.69 x 16.54	1 / 8	2	B	11 x 17	12 x 18	2	100 KiloBytes
A2	420 x 594	16.54 x 23.39	1 / 4	4	C	17 x 22	18 x 24	4	200 KiloBytes
A1	594 x 841	23.39 x 33.11	1 / 2	8	D	22 x 34	24 x 36	8	400 KiloBytes
A0	841 x 1189	33.11 x 46.81	1	16	E	34 x 44	36 x 48	16	800 KiloBytes
2A0	1189 x 1682	46.81 x 66.22	2	32					1.6 MegaBytes
					F	28 x 40	varies		600 KiloBytes
					G	11 x (22 1 / 2 to 90)	varies		
					H	28 x (44 to 143)	varies		
					J	34 x (55 to 176)	varies	Sizes G, H, J, and K	
					K	40 x (55 to 143)	varies	are US roll sizes	

Paper size references: MIL-M9868-D, Microfilming of Engineering Documents, 35MM, Requirements for, 10-1-70 and amendments 1 and 2, 2-12-82 and 9-20-82; MIL-STD-804B Format and Coding of Aperture, Copy and Tabulating Cards Engineering Data Micro-reproduction System, 15 August, 1966; ANSI Y 14.1, 1980, Drawing Sheet Size and Format, published by ASME (American Society of Manufacturing Engineers), New York; Metric standards first published in 1922 by DIN (Deutsches Institut für Normung) (German Institute for Standards) [http://www.DIN.de] Now used worldwide as ISO 216. (See page 6 for ISO references)

Newspapers: A double truck (center fold) full broadsheet is 24 in x 36 in, equivalent to an old D size drawing in size. Because a newspaper page would be scanned at a higher resolution and contains detailed graphics, a double truck would require about 1 MegaByte and a single full broadsheet page (18 by 24 inches) would require about 1 / 2 MegaByte. See also NAA (Newspaper Association of America) [http://www.NAA.org]

Scanned Microforms

1 roll of 16 mm microfilm (100 ft, ~30 meters) (24X reduction) = 2,500 letter size images = 1 box = 1 file cabinet drawer = 125 MegaBytes
1 roll of 35 mm microfilm (100 ft) (12X reduction, open spacing, normal scan) = 1,000 letter size images = 50 MegaBytes
1 microfiche (105 mm film) (24X reduction) = 100 letter size images = 5 MegaBytes (average); 200 microfiche = 20,000 images = 1 GigaByte

In many record series, each microfiche contains only a few images because each fiche represents a single record in the series (e.g. one fiche per person in a personnel record series). In this case filming breaks on records, rather than being continuous. To a lesser extent this is also true for roll film. In these cases, the amount of storage required depends on the number of images on the film, not the number of microfiche or the number of rolls of film. A full, standard 24X microfiche has 7 rows of 14 letter size (8 1/2 x 11 or A4) images for a total of 98 images.

As with any microform, scanned aperture card images require the same storage as images scanned from the paper original of the document in the aperture.

Compression

All documents are stored and transmitted in compressed format. All compression formats are assumed to be lossless or used with a lossless setting, except MPEG (Moving Picture Experts Group), unless otherwise stated. Lossless or non-destructive compression (as opposed to lossy or destructive compression) does not change the document. That is, a decompressed document is identical to the original document before compression was done. Lossless compression is often needed to meet legal requirements for document storage. The most common form of one bit (per pixel), bitonal (The two tones of color are two shades of grey which are black and white.), lossless compression, used in TIFF G4 and Adobe PDF (Portable Document Format), is the CCITT G4 (Group 4) facsimile compression format. Before using any other form of compression, it is often useful to evaluate the cost savings of moving to the less common format. The CCITT (Comité Consultatif International pour le Télégraphe et le Téléphone) (International Telegraph and Telephone Consultative Committee) is now a part of the ITU (International Telecommunications Union) [http://www.ITU.int]] The G4 ITU recommendation T.6 (11/88), Facsimile coding schemes and coding control functions for Group 4 facsimile apparatus, is on pages 48-57 of the CCITT Blue Book, Volume VII - Fascicle VII.3, Terminal Equipment and Protocols for Telematic Services, Recommendations T.0 - T.63, ISBN 92-61-03611-2

Archive Builders http://www.ArchiveBuilders.com for more details +1 (310) 937-7000 22009v111G

Digital Image Sizes

Computer storage requirements for various digitized document types

Scanned Miscellaneous Documents

1 check (2 sided) (remittance) = 50 KiloBytes per item, 25 KiloBytes (1 sided), less if no patterns are present.

1 credit card receipt (long: 3 1/4 x 7 7/16 inches, 2 sided) (remittance) = 35 KiloBytes, short (3 1/4 x 5 in., 2 sided) = 25 KBytes The long size credit card receipt is the same as an 80 column punch card, which was based on the older 90 column, round hole, punched card, which in 1890 was based on the size of the old US dollar bill (before 1929). US dollar bills are now 6.14 x 2.61 inches (~156 x ~66 mm), before 1929, US dollar bills were 7.4218 x 3.125 inches (~189 x ~79 mm). [~ emphasizes an approximation, rather than a precise measure.]

1 library book (average, scanned in black and white) = 10 MegaBytes; 50 books = 500 MBytes = 1 CD; 100 books = 1 GByte

Pages per Second (Internet Communications)

DSL (Digital Subscriber Line) = 1/2 to 8 Mbits (Megabits) per second = 1 to 15 pages per second (about ~ US$ 50.00 per month)

Modem = 56 Kbit (Kilobits) per second = 3 pages per minute (about ~ US$ 30.00 per month for a standard phone line) (2 bytes per baud (cycle))

ISDN (2 voice channels) = 128 Kbit per second = 10 pages per minute (~ US$ 100.00 per month) (ISDN charge)

Cable (TV) modem =~ 500 Kilobits per second = 1 page per second (about ~ US$ 50.00 per month)

T1 (24 voice channels, 64 Kilobits/sec each) = 1.544 Mbit (Megabit) per second = 3 pages per second (~ US$ 1,000.00 per month)

Ethernet (CSMA/CD) = 1 Mbit per second (effective) or 10 Mbit per second (nominal) = 2 pages per second

OC3 ATM (Optical Carrier, Asynchronous Transfer Mode) = 155 Mbit per second = 300 pages per second (1 1/2 books per second)

OC192 (SONET: Synchronous Optical NETwork fiber) = 10 Gbit per second = 20,000 pages (2 file cabinets) per second (1,000 books per second)

Dense Wavelength Division Multiplexing (DWDM) with 32 OC192 channels = 320 Gigabits per second = 64 file cabinets per second

Dense Wavelength Division Multiplexing (DWDM) with 80 OC1536 channels = 6.4 Terabits per second = 1,600 file cabinets per second
Announced by Nortel Networks On October 12, 1999: [http://www.NortelNetworks.com/corporate/news/newsreleases/1999d/10_12_9999633_80gigabit.html]

Optical carrier frequency (1,300 nm) (single-mode dark fiber) = 230 THz (TeraHertz) (about 2,000 cycles (baud) are used for every OC1536 bit transmitted)

Optical carrier frequency (1,550 nm) (coaxial dark fiber) = 193 THz (~1 Petabit per second per fiber at 1 byte per baud); see [http://www.Omni-Guide.com]

1 Petabit per second = ~2 billion pages per second = ~200 thousand file cabinets per second = ~10 million books per second (Intercity fiber cables often have 144 fibers)
See also ITU (International Telecommunications Union) [http://www.ITU.int] TIA (Telecommunications Industry Association) [http://www.TIAonline.org] The Internet Society (ISoc) [http://www.ISoc.org] The Internet Corporation for Assigned Names and Numbers (ICANN) [http://www.ICANN.org] The next Internet: [http://www.Internet2.edu]

(Office Color) Scanned Letter Size Pages (View-only images, 100 dpi, no OCR possible)

1 scanned page (100 dpi) (8 1/2 by 11 inches, A4) = 100 KiloBytes (KByte) (on average, office color, including grayscale, compressed)

1 file cabinet (4 drawer) (10,000 pages on average) = 1 GigaByte (GByte) = 2 CDs (ROM or WORM)

5 file cabinets = 1 DVD-R (WORM) (see below)

1,000 file cabinets = 1,000 GigaBytes = 1 TeraByte (TByte); 1,000 file cabinets = 200 DVDs

1 box (in inches: 12 wide x 15 long x 10 deep) (300 x 375 x 250 mm) (2,500 pages) = 1 file drawer = 2 linear feet (500 mm) of files = 250 MegaBytes

4 boxes = 8 linear feet = 1 file cabinets = 1 GigaByte; 4,000 boxes = 8,000 linear feet = 1,000 GigaBytes = 1 TeraByte

In general, when compressed, the digital files for document images scanned in office quality view-only-color (100 dpi) (no OCR possible) are about twice the size of document images scanned in a bi-tonal, black and white format, and then G4 compressed. In office quality color scanning, the scanned color differences aid users in reading a document and in increasing the quality of OCR (Optical Character Recognition) done at 150 dpi and higher resolutions. Office quality view-only-color scanning is generally at a lower resolution (100 dpi) than black and white scanning (300 dpi, required for bi-tonal OCR). Office quality grayscale-OCR-color (150 dpi) includes (has subsumed) the process of grayscale scanning which can increase OCR accuracy (at or above 150 dpi) when using low resolution scanning (lower than the 300 dpi generally required for bi-tonal OCR). Grayscale OCR is also called 3D OCR. (3 Dimensional OCR)

For the study of color (and color perception), see also CIE, the International Commission on Illumination (Commission Internationale de l'Eclairage) (Internationale Beleuchtungskommission) [http://members.eunet.at/CIE] See also SPIE, the International Society for (Photo) Optical Engineering, [http://www.SPIE.org]

(Office Color) (Includes Grayscale Scanning) Engineering Drawings / Aperture Cards

View-only-color (100 dpi, no OCR possible): 1 E size drawing (A0) (48 inches by 36 inches, with overscan) = 16 letter size pages (8 1/2 by 11 inches, A4) = 1,600 KiloBytes (1.6 MegaBytes) D size = 8 letter size pages; C size = 4 letter size pages; B size = 2 letter size pages; A size = 1 letter size page

Grayscale-OCR-color (150 dpi color, including a separate 300 dpi bi-tonal image): 1 E size drawing (48 inches by 36 inches) = 8 MegaBytes

Visually-unaltered-color (150 dpi color, including a separate 300 dpi bi-tonal image): 1 E size drawing (48 inches by 36 inches) = 32 MegaBytes

Raw-color (uncompressed) (grayscale only, 400 dpi): 1 E size drawing (48 inches by 36 inches) = 320 MegaBytes

Digitized Multimedia Formats

1 hour of compressed color video = 2 GigaBytes (DVD, MPEG 2) (image quality dependent) (On a DVD, 4 GigaBytes ~= One 2 hour feature length movie.)

1 hour of audio = 10 MegaBytes (dictation, answering machine) to 500 Mbytes (CD quality audio) (A CD holds 74 minutes of music.)

1 color picture = 10 KiloBytes (thumbnail) to 5 MBytes (for each of 100 photos on a 500 MByte photo CD)

The size of a compressed image file depends on the resolution (dpi: dots per inch) and the detail (information) in the photograph. The detail in a photograph is dependent on the size of the negative and the quality of the film and the camera and lens (It is not related to the print size unless the print is smaller than the negative). The resolution of the scan should be chosen to match the detail of the photograph. For most cameras, films, and formats 35mm and smaller, the 5 MByte Photo CD format (2048 by 3072 pixels) captures all the information in the image. Note that this is in dots per image rather than dots per inch. Displays are also given in dots per image (Horizontal x Vertical: e.g. 1280 x 1024), with the horizontal dimensional always being given first.

Archive Builders **http://www.ArchiveBuilders.com for more details** **+1 (310) 937-7000** **22009v111G**

Digital Image Sizes

DVD Digital Video Disc

(See also [http://www.DVDdemystified.com/dvdfaq.html])

DVD (commonly Digital Video Disc) (A DVD is the same physical size as a CD.) DVD stands for Digital Versatile Disc, by vote of the committee that controls the trademark DVD, the DVD Forum. [http://www.DVDForum.org] All capacities are given in commercial units: e.g.: 1 GigaByte = 1 Billion Bytes; 1 MegaByte = 1 Million Bytes

NB: When you calculate the amount of storage you will need on a given CD or DVD (using the table below), be sure that the units you are using for the size (amount) of data you plan to record are given in commercial rather than computer units. If you are not sure that the size (amount) of your data is given in commercial units, then add 10 (ten) percent to the size (amount) of data you plan to record. In all cases, you should leave yourself some headroom (of at least 5 percent) for last minute changes. (This can be reduced as you gain experience.) If, in addition to the normal headroom allowance, you are also uncertain of the (data size) units used, it is best to allow a total of 15 percent for headroom.

Disc Type	Acronym	Media Type	Side A Top Layer	Side A Bottom Layer	Side B Top Layer	Side B Bottom Layer	Total Storage Capacity
120 mm (4 3/4 inch) DVD	DVD-R** (SS)	DVD Recordable	4.70 GigaBytes	Not Available	Not Available	Not Available	4.70 GigaBytes
	DVD-R** (DS)	DVD Recordable	4.70 GigaBytes	Not Available	4.70 GigaBytes	Not Available	9.40 GigaBytes
	ROM (DS/DL)	Read Only Memory	4.27 GigaBytes	4.27 GigaBytes	4.27 GigaBytes	4.27 GigaBytes	17.08 GigaBytes
	RW & RAM	ReWriteable Random Access Memory	4.70 GigaBytes	Not Available	4.70 GigaBytes	Not Available	9.40 GigaBytes
80 mm (3 1/8 inch) DVD	DVD-R** (DS)	DVD Recordable	1.46 GigaBytes	Not Available	1.46 GigaBytes	Not Available	2.92 GigaBytes
	ROM (DS/DL)	Read Only Memory	1.33 GigaBytes	1.33 GigaBytes	1.33 GigaBytes	1.33 GigaBytes	5.32 GigaBytes
	RW & RAM	ReWriteable Random Access Memory	1.46 GigaBytes	Not Available	1.46 GigaBytes	Not Available	2.92 GigaBytes
HD-DVD Future: ~2006 120 mm	DVD-R** (DS)	DVD Recordable	16+ GigaBytes	Not Available	16+ GigaBytes	Not Available	32+ GigaBytes
	ROM (DS/DL)	Read Only Memory	16+ GigaBytes	16+ GigaBytes	16+ GigaBytes	16+ GigaBytes	64+ GigaBytes
120 mm CD	All (SS/SL)	All (SS/SL only)	682* MegaBytes	Not Available	Not Available	Not Available	682* MegaBytes
80 mm CD	All (SS/SL)	All (SS/SL only)	194 MegaBytes	Not Available	Not Available	Not Available	194 MegaBytes

SS (Single Sided), DS (Double Sided), SL (Single Layer), DL (Double Layer), SS/SL (Single Sided / Single Layer), DS/DL (Double Sided / Double Layer), DS/SL (Double Sided / Single Layer per side), DS/ML (Double Sided / Mixed Layer; one side 1 layer, other side 2 layer), HD (High Density); Top Layer (Layer 1), Bottom Layer (Layer 0)

* CD capacities have always been advertised as 650 MegaBytes using the older computer based MegaByte (1,048,576 Bytes) size. Using the new commercial standard units of 1 Million Bytes per MegaByte, a CD holds 682 MegaBytes. DVD capacities, however, are always stated in the new, smaller, commercial units.

** DVD-R (Recordable) and CD-R are the equivalent of WORM (Write Once, Read Many) The DVD-R capacity listed above, of 4.7 GigaBytes per side, is for discs and DVD writers that conform to the new DVD-R standard (DVD-R 2.0). The new DVD-R discs and writers are available now. The older (DVD-R 1.0) capacity of DVD-R discs is 3.95 GigaBytes per side for a total of 7.9 GigaBytes for a two sided disc.

DVD 16X

CD drives read at up to 40X speed. Music CDs are listened to at 1X. Listening to music at 2X or at an X other than 1X does not make sense, except for Alvin and the Chipmunks.

CD and DVD Xs do not always mean that the entire CD or DVD will be read at X times the normal speed. This is because CDs and DVDs are meant to be read at a Constant Linear Velocity (CLV), which means that a CD or DVD rotates faster when reading the shorter inner tracks. Some high speed readers do not increase their rotational speed when reading the inner tracks. The rotational speed on these readers is a Constant Angular Velocity (CAV) (one rotation sweeps out an angle of 360 degrees).

A good rule of thumb is to reduce the X speed by 25 percent. A 16X DVD drive would transfer an entire DVD movie at 12X fast forward speed and would therefore read about 4 GigaBytes in about 1 / 6 hour or 10 minutes at a data rate of about 25 GigaBytes per hour. Restoring a 250 GigaByte Database with 10 of the 16X DVD readers would require about 1 hour. In a short time the DVD drives should approach the CD drive cost of 50 US dollars each, so that 10 of the DVD drives would cost about 500 US dollars. Restoring a 2.5 TeraByte database with 10 of the 16X DVD drives would require 10 hours. A 10 TeraByte database would require 4 times as many drives (40) to restore the database in the same time (10 hours). When restoring files, some time is required to create catalog entries. For large files this is less of a problem.

DVD Multimedia

6 channel (theater quality surround sound) (5.1, Dolby AC-3) / 96 KHz audio / 24 bit audio, 8 language tracks, 32 subtitle tracks, and about 135 minutes (long enough to accommodate 94% of all movies) of high quality video (720 horizontal pixels) on each of 4 layers. DVDs support runtime editing so that all ratings of a movie are on the same DVD; 'R' rated scenes can be skipped, without interruption, as the DVD is played. The file format is ISO 13346 UDF (Universal Disc Format), which harmonizes all CD recording standards including ISO 9660. A future technology, 3rd generation blue lasers [sort of a blue light special, as blue light has a wavelength about half that of red light], should yield a 64+ GigaByte DVD ROM for HDTV. [N.B. Optical disc is spelled with a ‘c’ as in music disc. Magnetic disk is spelled with a ‘k’ as in harrow disk.] [For a DVD with a two layer side, to reduce inter-layer crosstalk, the minimum pit length of both layers is increased from .40 um to .44 um. This results in longer (and therefore fewer) pits for more effective reading of the data.] See also [http://www.DVDdemystified.com/dvdfaq.html]

DVD Audio

DVDs can be used to record audio only, with no video. In addition, DVD audio includes various still images. DVD audio is different than the audio that is used as part of DVD video.

The DVD audio standard is for up to 6 channels, a sampling rate of 48, 96, or 192 KHz, and a sample size of 16, 20, or 24 bits. With 24 bit samples taken at a 192 KHz rate, this provides a 96 KHz frequency response and a 144 dB dynamic range. DVD audio can also provide for a lossless audio compression of about 2 to 1 which would have a playing time of 120 to 140 minutes for two-channel 192 KHz / 24 bit recordings for a single layer. Each DVD disc can have up to 4 layers, 2 layers per side.

DVD audio includes various still image modes for synchronized lyrics, navigation, etc. DVD audio allows up to 16 still graphics per track (or slightly more, depending on the compression ratio) and a set of limited transitions.

The audio used in DVD video can also be used without the video. This produces a stereo, DVD quality, play time of over 55 hours at 192 Kilobits per second (compressed) for a single layer and over 200 hours for a 4 layer DVD disc. Lower quality sound can be recorded as computer files on a DVD for much longer play times. At a compressed audio rate of 16 Kilobits per second (in the low range of telephony quality), this is 9 million seconds, 150 thousand minutes, 2,500 hours, 100 days, 15 weeks, or 3 months of audio on a 4 layer DVD disc. (Each of the 24 T-1 telephony voice channels carries 64 Kilobits per second: 8 thousand 8 bit audio (sound or volume) samples per second.) See also AES (Audio Engineering Society) [http://www.AES.org]

Bit Fade and Copying

Like all storage and communications media, CD and DVD discs have the property that bits stored on them fade. Every day, some of the stored bits fade away. CDs and DVDs have an error correcting code (ECC) that can correct (replace) the lost bits. Eventually, there are too many lost bits to be corrected. This is the basis for the estimated lifetimes of CD and DVD media. Rather than an estimate, ANSI/AIIM (Association for Information and Image Management [AIIM.org]/American National Standards Institute [ANSI.org]) MS59-1996 media error monitoring and reporting standard, which complements the ANSI X3.131, media error hardware interface, provides a means of directly counting the number of bad bits (the raw error rate) on a given CD or DVD. This gives a disc-by-disc reading on when to copy the data on the disc, and indicates exactly which discs will actually last (protect the data for) the disc’s projected lifetime (up to 100 years). Until commercial, end user implementations of MS59 are available for checking discs, many users are following a practice of copying CDs and DVDs every five years, regardless of the nominal warranty period.

Archive Builders http://www.ArchiveBuilders.com for more details +1 (310) 937-7000 22009v111G

Digital Image Sizes

Computer storage requirements for various digitized document types

Pixels per Image

Most document imaging resolution measures are in pixels (PICture ELement) per inch (or per mm - millimeter), and are commonly referred to as dpi (dots per inch) or dpmm (dots per mm). Most motion picture and still-photographic resolution measures are in pixels per image. This is most commonly seen in the 525 lines of NTSC (National Television System Committee), 625 lines for PAL (Phase Alternating Line) and SECAM (Sequential Couleur Avec Memoire or Sequential Colour with Memory), resolution of television images. No matter how physically large or small an NTSC television image is displayed, there are only 525 lines of vertical resolution (480 viewable). The computer equivalent of this is 640 by 480 pixels in a standard computer image. In pixels per image the horizontal resolution is given first. If the horizontal dimension is larger than the vertical dimension in pixels, the image or display is said to be landscape, if the horizontal is smaller, the image or display is said to be portrait. See also SMPTE (Society of Motion Picture and Television Engineers) [http://www.SMPTE.org]

The physical size of the display is an important element in the design of a document imaging workstation. A 20 or 21 inch nominal diagonal size, or exactly a 20 ± 1/4 inch VIS (Viewable Image Size) is most commonly used for CRTs (Cathode Ray Tubes), with an equivalent 18 ± 1/4 inch VIS being the most common for flat panel displays. These sizes, or larger, are especially important for extended use, or the accommodation of viewers who use bifocal glasses.

Pixels per Image: Video Image Resolutions

Computer screen resolutions are chosen to have an aspect ratio (the ratio of width to height) of 4 to 3 (the 'golden ratio' of the art world) and to have the number of pixels be an integer multiple of a power of 2. (Powers of 2 are given here as 2**N for the Nth power of 2). When a prefix is added to the word pixel it can be shortened to pel (Picture ELement). A 1 million pixel display is then a 1 MegaPel display. See also IEEE (The Institute of Electrical and Electronics Engineers) [http://www.IEEE.org] ACM (Association for Computing Machinery) [http://www.ACM.org] NAB (National Association of Broadcasters) [http://www.NAB.org]

IBM PC: CGA (Color Graphics Adapter) 320 x 200, EGA (Enhanced Graphics Adapter) 640 x 350, VGA (Video Graphics Array) 640 x 480.

IBM PC compatible: VGA 640 x 480 (This is the standard default screen resolution when a display card is reset to troubleshoot a problem with the display.), SVGA (Super VGA) or XGA (eXtended Graphics Array) 800 x 600, XVGA (eXtended VGA) 1024 x 768; and SXGA (Super XGA) or UVGA (Ultra VGA) 1280 x 1024 (although SVGA, SXGA, XVGA, and UVGA can mean anything that is more than the VGA's 640 x 480), UXGA (Ultra eXtended Graphics Adapter) is often 1600 x 1200. Because the meaning of IBM PC compatible acronyms is a marketing decision, no absolute meaning should be imputed to them.

The DVD NTSC resolution is 720 x 480 and the DVD PAL/SECAM resolution is 720 x 576. (Twentieth Century commercial television)

Common Display Resolutions

(An aspect ratio of [a x b] [c x d] is equal to [a * c x b * d] when expressed using matrix arithmetic.)

In all cases, the actual numeric resolutions should be used in place of an acronym. The acronyms are given here for assistance in interpreting text that does not include a numeric resolution. An acronym can be used following the numeric resolution, for reference, but the acronym may lead to extended discussions.

640 x 480	pixel resolution (VGA standard computer screen resolution) = [2**5 x 2**5] [4 x 3] [5 x 5]
800 x 600	(usually XGA, sometimes SVGA) = [2**3 x 2**3] [4 x 3] [25 x 25]
1024 x 768	(often XVGA, less often UVGA) = [2**8 x 2**8] [4 x 3]
1152 x 900	(Sun Microsystems) 1152 x 870 (Mac) (1152 = 2**4 x 72 typeset points per inch). Some Sun Microsystems and Apple / Mac screen resolutions were chosen so that the actual screen resolutions were 72 dpi to match the 72 points per inch used in typesetting.
1280 x 1024	(more often SXGA, sometimes UVGA, less often XVGA) = [2**8 x 2**8] [5 x 4]
1600 x 1200	(often UXGA) (high resolution document imaging workstation) = [2**4 x 2**4] [4 x 3] [25 x 25]
1920 x 1200	(HDTV) The computer version of HDTV (High Definition TV) resolution is 1920 x 1200 ([Sun.com] Microsystems) and has the HDTV 16 to 9 aspect ratio. The 1920 x 1200 resolution is designed to match the NTSC derived HDTV video resolutions of 1920 x 1080 and old analog HDTV (NTSC derived) resolution of 1920 x 1035 and the PAL and SECAM derived analog HDTV video resolution of 1920 x 1152 (1152 = 2 x 576). The current standard HDTV resolutions are 1280 x 720 and 1920 x 1080. The actual resolution of HDTV streams transmitted will usually be 1920 x 1088, because MPEG-2 requires the number of lines to be in multiples of 16 (1088 lines = 68 x 16).
1800 x 1440	(very high resolution grayscale document imaging workstation) = [72 x 72] [25 x 20]
2048 x 1536	(very high resolution grayscale document imaging workstation) = [2**9 x 2**9] [4 x 3]

Pixels per Image: Single Frame (Still) Image Resolutions

See also [http://www.Kodak.com]

The Kodak PhotoCD family of resolutions: (Based on a 2 x 3 portrait aspect ration and an integer power of 2. The multiple of the base gives the number of pixels per image relative to the base image size in pixels.) A Kodak PhotoCD contains five resolutions for each image: 1/16 Base through 16 Base. (The average compressed file size containing all five resolutions is about 5 MegaBytes per image.) A Kodak Pro PhotoCD contains the five resolutions for each image found on a PhotoCD plus a sixth 64 Base resolution. PhotoCD images are intended for true color, continuous tone images. 64 base Kodak Pro PhotoCD scanning does not always provide adequate resolution for 35 mm aperture card images (monotone microform images with a steep gamma curve) (nominally bi-tonal or black and white).

1/16	base (thumbnail, index print on CD cover)	.024576 megapixel image	=	128 x 192	[2 x 3] [2** 6 x 2 ** 6]
1/4	base (largest Kodak size that is smaller than 480 x 640 for display on TV)	.098304 megapixel image	=	256 x 384	[2 x 3] [2** 7 x 2 ** 7]
1	base	.393216 megapixel image	=	512 x 768	[2 x 3] [2** 8 x 2 ** 8]
4	base (largest Kodak size that is smaller than 1920 x 1152 for HDTV)	1.572864 megapixel image	=	1024 x 1536	[2 x 3] [2** 9 x 2 ** 9]
16	base (captures all the resolution on most 35 mm film images)	6.291456 megapixel image	=	2048 x 3072	[2 x 3] [2**10 x 2 **10]
64	base (captures all the resolution for most film formats larger than 35 mm)	25.165824 megapixel image	=	4096 x 6144	[2 x 3] [2**11 x 2 **11]

Archive Builders http://www.ArchiveBuilders.com for more details +1 (310) 937-7000 22009v111G

Digital Image Sizes

Computer storage requirements for various digitized document types

Medical Records

1 Chest X-ray (14 x 17 inches) = 1 MegaByte: 150 dpi (dots per inch), 12 bits (compressed) (Wavelet compression, lossless mode, has FDA 510(k) approval.) (12 bits per pixel provide 4,096 shades of gray.) 150 dpi, 12 bit images are recommended by the American College of Radiology for primary reads. See also ACR (American College of Radiology) [http://www.acr.org] RSNA (Radiological Society of North America) [http://www.RSNA.org] FDA (United States Food and Drug Administration) [http://www.FDA.gov] HIMSS (Health Information Management Systems Society) [http://www.HIMSS.org]

A lossy compression, 14 x 17 Chest X-ray = 200 KiloBytes (For secondary reads: wavelet compression, lossy mode, has FDA 510(k) approval.)

X-rays that are originally recorded digitally rather than on film provide a resolution (image depth) of 16 bits per pixel which records 65,536 shades of gray per pixel. More shades of gray allow doctors to see very fine variations in the health of tissues, increasing the early detection of disease.

Aerial Photography, Digital Orthophotography, and Remote Sensing Pixel Sizes

Aerial photography uses photographs taken from the air, recording the visible electromagnetic spectrum (light), as maps of geographic areas. Remote sensing includes photographs taken from the air and from beyond the atmosphere of areas on the earth and other celestial bodies, using many segments of the electromagnetic spectrum including visible light, ultraviolet, infrared, and radar illumination. Digital orthophotography digitally rectifies the pixels of digitized aerial photographs into a continuous map, usually registered to a layer of a GIS (Geographic Information System).

For cities, 2 inch to 6 inch pixels are popular for digital orthophotography. A digital orthophotograph of a 500 square mile city using 6 inch pixels would have 4 pixels per square foot, 100 million pixels per square mile (There are approximately 25 million square feet per square mile.), for a total of 50 GigaPels (50 billion pixels). Using 8 bit uncompressed grayscale or using 24 bit color with an estimated lossless three-to-one compression, this digital orthophographic image would require 50 GigaBytes to store. If 2 inch pixels were used, a 500 square mile city would have 9 times as many pixels or 450 GigaPels requiring 450 GigaBytes to store using the same assumptions. Using 2 inch pixels a 50 square mile city would have 45 GigaPels requiring 45 GigaBytes to store using the same compression assumptions. The metric equivalents are 50 millimeter (mm) and 100 mm pixels, which are respectively 400 and 100 to the square meter. For a 1 thousand square kilometer city this would be 100 GigaPels using 100 mm pixels and would require 100 GigaBytes to store. Using 50 mm pixels for a 1 thousand square kilometer city, this would require 400 GigaPels requiring 400 GigaBytes to store. A 100 square Kilometer city, using 50 mm pixels would be imaged in 40 GigaPels, which would require 40 GigaBytes to store.

In digital orthophotography, in addition to color, each pixel has an associated z axis value, the height of the pixel above sea level. When added to the x and y Cartesian coordinates of the pixel, the z values construct a digital terrain model over which the image can be mapped as a surface. This is similar to the way that images are created in virtual reality. By adding a t value, a 4 fourth dimension that represents a specific point in time, animations can be done telling a geologic story or the developmental history of a city.

In remote sensing (satellite imagery such as weather photographs or images for crop quality assessment or storm damage / flooding), a 24 bit color image of an area 1 thousand kilometers by 1 thousand kilometers, using 100 meter pixels (pixels that are 100 meters by 100 meters), would contain 100 million pixels. Estimating a lossless three-to-one compression this would require 100 MegaBytes to store. The pixels used can be of any size. In astronomy, a single pixel can include an entire earth type planet (10 thousand kilometer pixels = 10 Mm, 10 MegaMeter pixel), a sun type star (1 million Kilometer pixels = 1 Gm, 1 Gigameter pixel), or a galaxy (100 thousand light year pixels =~ 1 Zm, 1 Zettameter pixel). The largest practical pixel is a 400 Ym, 400 Yottameter pixel, the diameter of the observable universe.

See also ACSM (American Congress on Surveying and Mapping) [http://www.SurvMap.org] ASPRS (American Society for Photogrammetry and Remote Sensing) [http://www.ASPRS.org] IAU (International Astronomical Union) [http://www.IAU.org]

GIS (Geographic Information System) Data Storage Requirements

GIS data, average, in city: 1 square mile = 50 MegaBytes; 20 sq. miles = 1 GigaByte, 1 square Kilometer = 20 MegaBytes, 50 sq. Kilometers = 1 GigaByte

Digital Orthophoto data, 6 inch pixels, uncompressed monochrome (or losslessly compressed color), 1 square mile = 100 MegaBytes; 10 square miles = 1 GigaByte; Using 100 mm pixels: 1 square meter = 100 Bytes, 1 square Kilometer = 100 MegaBytes, 10 square Kilometers = 1 GigaByte

Digital Orthophoto data, 2 inch pixels, uncompressed monochrome (or losslessly compressed color): 1 square mile = 1 GigaByte; Using 50 mm pixels: 1 square meter = 400 Bytes, 1 square Kilometer = 400 MegaBytes, 2.5 square Kilometers = 1 GigaByte

See also URISA (Urban and Regional Information Systems Association) [http://www.URISA.org]

Microprocessor and RAM (Random Access Memory) Design Rules (Pixel Size)

Semiconductors are made using digital photographic techniques (pixels). Recently, microprocessor production processes were improved from .25 micron (.25 um, micrometer) (250 nm, nanometer) design rules to .18 um (180 nm) design rules. This means that the pixel size for semiconductor devices is now slightly less than 1/5 micron (200 nm). A micron is one 1 millionth of a meter.

Using 200 nanometer (nm) pixels and assuming 1/25th of the area was used for active transistors, a 1 millimeter (mm) square area (about the size of the head of a pin) could hold 25 MegaPels (25 million pixels) and 1 million transistors. This is the basis for smart dust technology, developed at the University of California at Berkley, in which remote robots called motes could be built on ultra thin 1 millimeter square chips of silicon that float through the air and communicate with micro-lasers and micro-mirrors. Motes can remain suspended in the air for many hours, just like a cloud of dust or windblown seeds, collecting very detailed data. See also [http://robotics.eecs.Berkeley.edu/~pister/SmartDust] and International SEMATECH (SEmiconductor MAnufacturing TECHnology association) [http://www.SEMATECH.org]

The smallest practical pixel would be a pixel used as part of a halftone dot that represented the edge of the path of a sub-atomic particle, such as a neutrino. To create a smooth path in a specific color, a printed resolution of 2540 dpi (100 dpmm) would be used. Assuming a 1 ym (yoktometer) wide path, rendered as a 10 mm wide path, the width represented by each pixel would be 1/1 thousand ym. For a superstring (2 x 10**-35 m wide), the pixel width would be 20/1 trillion ym.

Halftone dots vary in size, but their centers are on a regular grid. Halftone dots are laser printed as an array of pixels. A 16 by 16 pixel array (or macropel) can represent any one of 256 shades of gray. As the number of black pixels printed in the center of the macropel array increases, the diameter of the halftone dot increases, creating the impression of a darker gray image. For this reason, a scanner that scans 8 bits (256 shades of gray) can be said to require 256 pixels (arranged in a 16 x 16 pixel array) to reproduce each scanned pixel as a halftone dot. This is the reason a 300 dpi scanner can be represented as 4800 dpi in advertisements (4800 dpi = 300 dpi x 16).

See also GATF (Graphic Arts Technical Foundation) [http://www.GATF.lm.com]

Digital Image Sizes

Units of Measure (Digital and Metric)

1 Byte (B) is defined as the set of bits used to represent 1 character. Commonly: 1 Byte (B) = 8 bits (b). (Byte & bit are best spelled out.) 8 bits can represent 256 different characters. ASCII (ANSI (American National Standards Institute) Standard Code for Information Interchange) (see [http://www.ANSI.org]) uses an 8 bit code to represent 256 characters. 1 ASCII Byte = 8 bits = 1 character 16 bits can be used to represent 65,536 different characters. Unicode uses a 16 bit code to represent 65,536 different characters (some of which are unassigned) to include most of the world's languages in the same, consistent character set. 1 Unicode Byte = 16 bits = 1 character See [http://www.Unicode.org] See also ISO (International Organization for Standardization) for the universal system of measurement known as SI (Système International d'unités). ("ISO" is not an acronym. "ISO" is a word, derived from the Greek isos, meaning "equal", which is the root of the prefix "iso-" that occurs in a host of terms, such as "isometric" (of equal measure or dimensions) and "isonomy" (equality of laws, or of people before the law).) [http://www.ISO.ch]

1 Hertz = 1 cycle per second (e.g. 1 clock cycle in a computer which corresponds roughly to the time required to execute 1 computer instruction. In these terms, a 1 GigaHertz computer executes 1 Billion instructions per second.). A 1,000 cycle per second signal or action is called a 1 KiloHertz signal or action (a 1 KHz signal), each cycle of such a signal is 1 millisecond (ms) long. See BIPM (Bureau International des Poids et Mesures) [http://www.BIPM.fr/enus] for metric units.

1 KiloByte	= 1,000 Bytes	= 1 Thousand Bytes	(KByte)			
1 MegaByte	= 1,000 KBytes	= 1 Million Bytes	(MByte)			
1 GigaByte	= 1,000 MBytes	= 1 Billion Bytes	(GByte)	= 1 Million KiloBytes		
1 TeraByte	= 1,000 GBytes	= 1 Trillion Bytes	(TByte)	= 1 Million MegaBytes	= 1 Billion KiloBytes	
1 PetaByte	= 1,000 TBytes	= 1 Quadrillion Bytes	(PByte)	= 1 Million GigaBytes	= 1 Billion MegaBytes	= 1 Trillion KiloBytes
1 ExaByte	= 1,000 PBytes	= 1 Quintillion Bytes	(EByte)	= 1 Million TeraBytes	= 1 Billion GigaBytes	= 1 Trillion MegaBytes
1 ZettaByte	= 1,000 EBytes	= 1 Sextillion Bytes	(ZByte)	= 1 Million PetaBytes	= 1 Billion TeraBytes	= 1 Trillion GigaBytes
1 YottaByte	= 1,000 ZBytes	= 1 Septillion Bytes	(YByte)	= 1 Million ExaBytes	= 1 Billion PetaBytes	= 1 Trillion TeraBytes

1 KiloHertz	= 1,000 Hertz	= 1 Thousand Hertz	(kHz)	10**+3	1 Kilometer	= 1,000 meters	= 1 Thousand meters	(km)
1 MegaHertz	= 1,000 KHertz	= 1 Million Hertz	(MHz)	10**+6	1 Megameter	= 1,000 Kmeters	= 1 Million meters	(Mm)
1 GigaHertz	= 1,000 MHertz	= 1 Billion Hertz	(GHz)	10**+9	1 Gigameter	= 1,000 Mmeters	= 1 Billion meters	(Gm)
1 TeraHertz	= 1,000 GHertz	= 1 Trillion Hertz	(THz)	10**+12	1 Terameter	= 1,000 Gmeters	= 1 Trillion meters	(Tm)
1 PetaHertz	= 1,000 THertz	= 1 Quadrillion Hertz	(PHz)	10**+15	1 Petameter	= 1,000 Tmeters	= 1 Quadrillion meters	(Pm)
1 ExaHertz	= 1,000 PHertz	= 1 Quintillion Hertz	(EHz)	10**+18	1 Exameter	= 1,000 Pmeters	= 1 Quintillion meters	(Em)
1 ZettaHertz	= 1,000 EHertz	= 1 Sextillion Hertz	(ZHz)	10**+21	1 Zettameter	= 1,000 Emeters	= 1 Sextillion meters	(Zm)
1 YottaHertz	= 1,000 ZHertz	= 1 Septillion Hertz	(YHz)	10**+24	1 Yottameter	= 1,000 Zmeters	= 1 Septillion meters	(Ym)
1 millisecond	= 1 / 1,000 second	= 1 Thousandth second	(ms)	10**-3	1 millimeter	= 1 / 1,000 meter	= 1 Thousandth meter	(mm)
1 microsecond	= 1 / 1,000 millisecond	= 1 Millionth second	(us)	10**-6	1 micrometer	= 1 / 1,000 millimeter	= 1 Millionth meter	(um)
1 nanosecond	= 1 / 1,000 microsecond	= 1 Billionth second	(ns)	10**-9	1 nanometer	= 1 / 1,000 micrometer	= 1 Billionth meter	(nm)
1 picosecond	= 1 / 1,000 nanosecond	= 1 Trillionth second	(ps)	10**-12	1 picometer	= 1 / 1,000 nanometer	= 1 Trillionth meter	(pm)
1 femtosecond	= 1 / 1,000 picosecond	= 1 Quadrillionth second	(fs)	10**-15	1 femtometer	= 1 / 1,000 picometer	= 1 Quadrillionth meter	(fm)
1 attosecond	= 1 / 1,000 femtosecond	= 1 Quintillionth second	(as)	10**-18	1 attometer	= 1 / 1,000 femtometer	= 1 Quintillionth meter	(am)
1 zeptosecond	= 1 / 1,000 attosecond	= 1 Sextillionth second	(zs)	10**-21	1 zeptometer	= 1 / 1,000 attometer	= 1 Sextillionth meter	(zm)
1 yoktosecond	= 1 / 1,000 zeptosecond	= 1 Septillionth second	(ys)	10**-24	1 yoktometer	= 1 / 1,000 zeptometer	= 1 Septillionth meter	(ym)

In the abbreviation for 1 microsecond (us), u is substituted for the symbol for the Greek letter mu.

Because light travels about 300 MegaMeters (Mm) in 1 second and has a wavelength of about 400 nm for blue light (about 700 nm for red light), the frequency of light is about 750 THz for blue light, about 430 THz for red light, and about 230 THz for the 1,300 nm light used in fiber optics). This is because speed (e.g.: C, the speed of light, which is a constant) = wavelength X frequency.

1,000 Bytes = 1 KiloByte (exactly 1 Thousand Bytes in common and legal usage) (exactly 1,024 Bytes = 2**10 = 2 to the 10th power in computer terms); 1,000 KBytes = 1 MegaByte (exactly 1 Million Bytes in common and legal usage) (exactly 1,024 KBytes = 1,048,576 Bytes = 2**20 = 2 to the 20th power in computer terms);

For marketing purposes, a given disk can hold more of the smaller commercial units than the larger computer units. For example a disk that contains 770 computer based MegaBytes (1,048,576 Bytes) sounds smaller than a disk that contains 807 of the commercial MegaBytes (1,000,000 Bytes), even though both disks hold exactly the same number of bytes of data. For both marketing purposes, and because of concern about lawsuits, only the commercial terms have been used in commercial descriptions in recent years.

Conversion from computer based terms to commercial terms. (Including the percent by which the computer terms are larger than the corresponding commercial terms.)

1,024 Commercial Bytes	= 1 Computer Based KiloByte	(a difference of 2.4 percent)	1,073,741,824 Commercial Bytes	= 1 Computer Based GigaByte	(a difference of 7.4 percent)
1,048,576 Commercial Bytes	= 1 Computer Based MegaByte	(a difference of 4.9 percent)	1,099,511,627,776 Commercial Bytes	= 1 Computer Based TeraByte	(a difference of 10 percent)

There are 1,048 Commercial KiloBytes in a Computer Based MegaByte, but only 1,024 Computer KiloBytes in a Computer Based MegaByte.

Computer units are given in powers of two (2**N) because the address space (size of memory, memory capacity) of a computer is determined by the number (N) of address lines available. A 32 bit computer has 32 address lines, has a 32 bit address space, and can address 2**32 (= 4,294,967,296) Bytes of RAM (Random Access Memory). The capacity of a disk or disc is determined by the number of sectors, tracks, platters, layers, and/or sides. These numbers are not based on powers of 2.

Paper, Trees, COLD Pages, Word Processor Pages, OCR Output Pages, and Commercial Scanning

1 pulp tree (loblolly pine) = 1/10th cord of wood = 10,000 pages = 1 file cabinet = 4 boxes = 1/2 GigaByte = 1 CD // 1 lumber tree (20 inch (500 mm) diameter, 110 ft (35 m) tall, 50 years old) = 1 cord = 10 pulp trees (8 in. (200 mm) diameter, 50 ft (15 meters) tall, 20 yrs old) = 1 cord = 4 ft x 4 ft x 8 ft = 128 cubic feet (3.5 cubic meters) as stacked for storage (75 cubic feet of wood, 2 cubic meters of wood) = 100,000 pages = 5 GigaBytes See also AFPA (American Forest & Paper Association) [http://www.AFandPA.org]

Microsoft Office 2000 (and later) documents: 20 KiloBytes per document for the XML repository + 5 KiloBytes per page //// 1 wordprocessor page, 1 office suite page (without the Office 2000 XML repository), or 1 OCRed (Optical Character Recognition) page = 5 KBytes (all pages listed above this section are scanned pages) For SGML (Structured Generalized Markup Language), HTML (HyperText ML), XML (Extensible ML), and CGM (Computer Graphics Metafile): see OASIS (Organization for the Advancement of Structured Information Standards) [http://www.OASIS-open.org] W3C (World Wide Web Consortium) [http://www.w3.org/XML]

1 compressed page of COLD (Computer Output to Laser Disc) or COOL (Computer Output On-Line) (including index) = 2 KiloBytes for letter size statements, 4 KiloBytes for 11 x 14 inch fanfolded greenbar computer sheet, 10 KiloBytes for All Points Addressable (APA) pages such as IBM AFP (Advanced Function Printing) and Xerox Metacode. For printing, see also Xplor International [http://www.Xplor.org]

Minimum commercial scanning cost for backfile conversion (more than 1 million pages) = about ~ 5 US cents per page

Compiled by Steve Gilheany, CRM, CDIA

Archive Builders http://www.ArchiveBuilders.com for more details +1 (310) 937-7000 **22009v111G**

Evolution of Intel Microprocessors: 1971 to 2003

Family	Trade Name (Code Name for Future Chips)	Clock Frequency in MegaHertz***	Millions of Instructions per Second	Date of Introduction	Number of Transistors	Design Rule (Pixel Size)	Address Bus Bits
80786	(Northwood)	3,000.0 MHz	TBA	2003	TBA	0.13 micron	64 bit
80786	(Madison)****	TBA	TBA	2003	TBA	0.13 micron	64 bit
80786	(Deerfield)****	TBA	TBA	2002Q2	TBA	0.13 micron	64 bit
80786	(McKinley)**	1,000.0 MHz	TBA	2001Q4	TBA	0.18 micron	64 bit
80786	Itanium**	800.0 MHz	TBA	2001Q1	TBA	0.18 micron	64 bit
80686	Pentium 4*****	1,500.0 MHz	*1,500.00 MIPS	2000Q4	TBA	0.18 micron	32 bit
80686	Pentium III	1,000.0 MHz	*1,000.00 MIPS	March 1, 2000	28.1 million	0.18 micron	32 bit
80686	P III Xeon	733.0 MHz	*733.00 MIPS	October 25, 1999	28.1 million	0.18 micron	32 bit
80686	Mobile P II	400.0 MHz	*400.00 MIPS	June 14, 1999	27.4 million	0.18 micron	32 bit
80686	P III Xeon	550.0 MHz	*550.00 MIPS	March 17, 1999	9.5 million	0.25 micron	32 bit
80686	Pentium III	500.0 MHz	*500.00 MIPS	February 26, 1999	9.5 million	0.25 micron	32 bit
80686	P II Xeon	400.0 MHz	*400.00 MIPS	June 29, 1998	7.5 million	0.25 micron	32 bit
80686	Pentium II	333.0 MHz	*333.00 MIPS	January 26, 1998	7.5 million	0.25 micron	32 bit
80686	Pentium II	300.0 MHz	*300.00 MIPS	May 7, 1997	7.5 million	0.35 micron	32 bit
80586	Pentium Pro	200.0 MHz	*200.00 MIPS	November 1, 1995	5.5 million	0.35 micron	32 bit
90586	Pentium	133.0 MHz	*133.00 MIPS	June 1995	3.3 million	0.35 micron	32 bit
80586	Pentium	90.0 MHz	*90.00 MIPS	March 7, 1994	3.2 million	0.60 micron	32 bit
80586	Pentium	60.0 MHz	*60.00 MIPS	March 22, 1993	3.1 million	0.80 micron	32 bit
80486	80486 DX2	50.0 MHz	*50.00 MIPS	March 3, 1992	1.2 million	0.80 micron	32 bit
80486	486 DX	25.0 MHz	20.00 MIPS	April 10, 1989	1.2 million	1.00 micron	32 bit
80386	386 DX	16.0 MHz	5.00 MIPS	October 17, 1985	275,000	1.50 micron	16 bit
80286	80286	6.0 MHz	0.90 MIPS	February 1982	134,000	1.50 micron	16 bit
8086	8086	5.0 MHz	0.33 MIPS	June 8, 1978	29,000	3.00 micron	16 bit
8080	8080	2.0 MHz	0.64 MIPS	April 1974	6,000	6.00 micron	8 bit
8008	8008	.2 MHz	0.06 MIPS	April 1972	3,500	10.00 micron	8 bit
4004	4004	.1 MHz	0.06 MIPS	November 15, 1971	2,300	10.00 micron	4 bit

* Approximately one instruction per processor clock cycle

** Itanium, formerly codenamed Merced, may be replaced by McKinley if further delayed.

*** 1 Hz (Hertz) = 1 cycle per second, 1 KHz (KiloHertz) = 1 thousand cycles per second; 1 MHz (MegaHertz) = 1 million cycles per second, 1 MegaHertz = 1 thousand KiloHertz, 1 GHz (GigaHertz) = 1 billion cycles per second; 1 GigaHertz = 1 thousand MegaHertz

**** .Deerfield is a low cost version of Madison.

*****.The Pentium 4 was formerly code named Willamette

TBA To be announced, Trademarks are the property of their respective holders.

[http://www.esc-ca.com/processors/intel/future.htm] (one source of data for future microprocessors)

[http://www.Intel.com/pressroom/kits/processors/quickref.htm] (source of data for released microprocessors)

MHz (MegaHertz) (Millions of processor cycles per second) The number of times the processor goes through one cycle. The start of a processor cycle is determined by a pulse (tick) from the processor's clock.

GHz (GigaHertz) (Billions of processors cycles per second). 1 thousand MHz = 1 GHz

MIPS: (Millions of Instructions per Second) with the introduction of the 80486 DX2, parallel instruction execution increased the number of instructions executed per processor cycle to approximately one instruction per cycle. Parallel instruction execution requires many more transistors, so the increase in the number of transistors has increased the number of instructions that can be executed per second faster than the clock cycle speed has increased. A larger transistor budget allows the addition of specialized instructions, which increase the microprocessor's speed in processing specialized information such as graphics by increasing the amount of information processed per instruction.

GIPS (GigaInstructions per Second) Billions of instructions per second. 1 thousand MIPS = 1 GIP

Design Rule: because the wires and components, including transistors, on chips are drawn photographically, the pixel size of the imaging process determines the width of the wires and the size of the transistors. The size of the transistors determines how many will fit on a chip of a given size. (The optimal size of a chip depends on the chip manufacturing processes. In general, chip size increases slowly over time.) The smaller the transistors, the more will fit on the chip, determining the chip's transistor budget. The size of the transistors also determines the transistor's switching speed. Smaller transistors switch faster. One micron is one one-millionth of a meter or about 40 millionths of one inch. Finally, the power required to switch smaller transistors is less, so smaller pixels in the design rules allow the batteries in laptop computers to last longer.

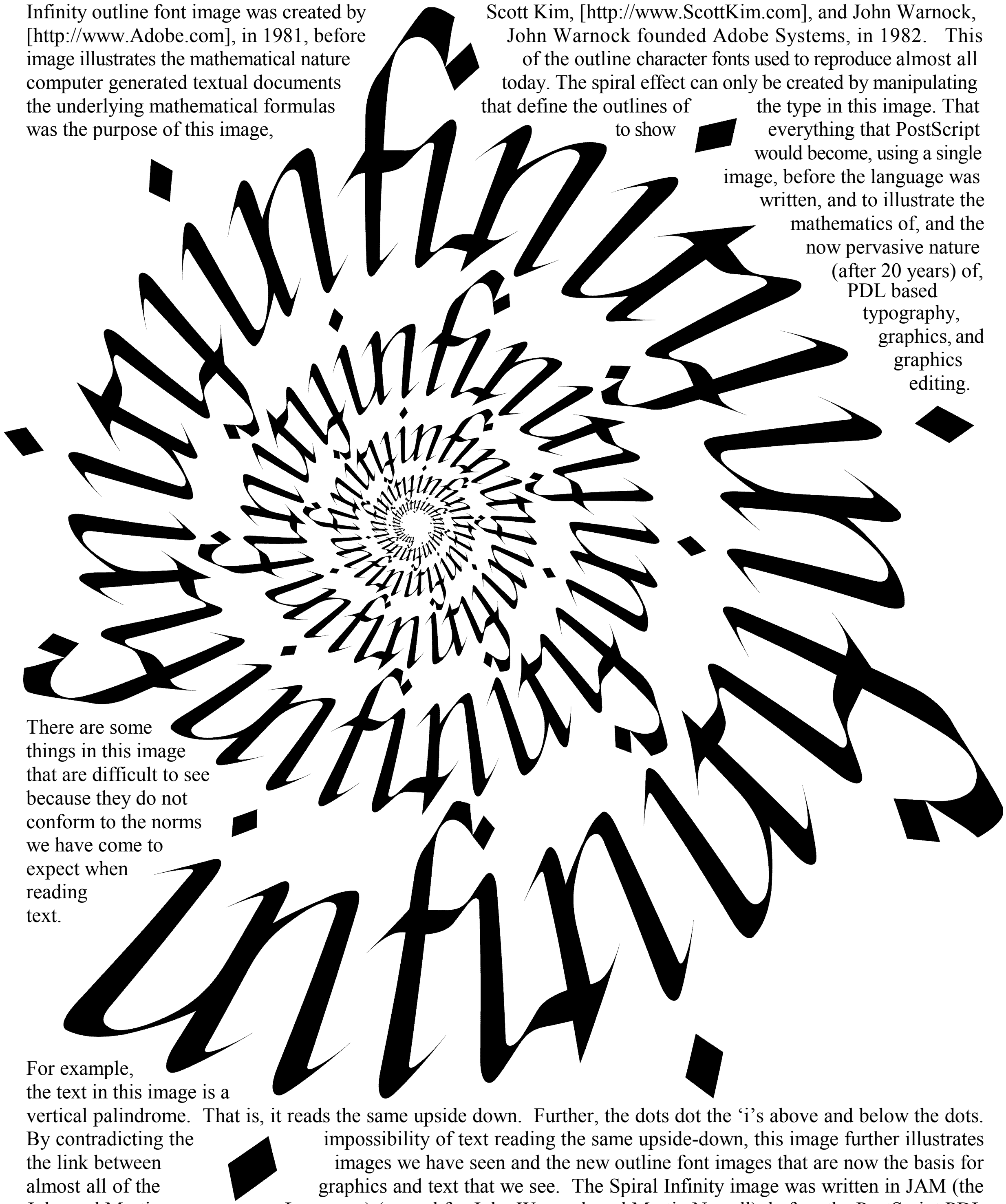
Number of Transistors: The number of transistors increases as the square of decrease in design rule size. Each reduction in design rule size is chosen to about double the number of available transistors (the transistor budget). [For example: (.25micron / .18 micron) x (.25 / .18) = 2.] The gradual increase in die size (the size of the chip) also increases the number of transistors.

Address Bus Bits: The address bus width in bits is based on the microprocessor chip family. (In the later chips of the 80686 family, some changes have been made to make more memory addressable under special circumstances, by using 36 bits to address 16 times as much memory as is possible with 32 address bits, but the generalized addressing structure is still 32 bits.) Each time a bit is added to the address bus width, the amount of memory (RAM: Random Access Memory) that can be addressed is doubled. 4 bit addresses allow the addressing of 16 bytes of memory (and extra work is necessary to address 256 bytes of memory). 8 bits allow the addressing of 256 bytes of memory (and extra work is necessary to address 65,536 bytes of memory). 16 bits can address 65,536 bytes of memory. 32 bits can address 4,294,967,296 bytes of memory (about 4 billion bytes). As memory prices drop, it becomes necessary to address over 4 billion bytes of memory. The 80786 family is due out from Intel in the year 2001. It will have a 64 bit address bus and will be able to address over 16 billion billion (16 quintillion) bytes of memory.

See also: [http://www.intel.com/intel/intelis/museum] Intel's history of the microprocessor.

Figure 25.

Spiral Infinity is the exemplar for the Interpress (Xerox), of the PostScript (Adobe), and later of the Acrobat PDF (Portable Document Format) (Adobe) outline font based PDLs (Page Description Languages). The Spiral Infinity outline font image was created by Scott Kim, [http://www.ScottKim.com], and John Warnock, [http://www.Adobe.com], in 1981, before John Warnock founded Adobe Systems, in 1982. This image illustrates the mathematical nature of the outline character fonts used to reproduce almost all computer generated textual documents today. The spiral effect can only be created by manipulating the underlying mathematical formulas that define the outlines of the type in this image. That was the purpose of this image, to show everything that PostScript would become, using a single image, before the language was written, and to illustrate the mathematics of, and the now pervasive nature (after 20 years) of, PDL based typography, graphics, and graphics editing.



There are some things in this image that are difficult to see because they do not conform to the norms we have come to expect when reading text.

For example, the text in this image is a vertical palindrome. That is, it reads the same upside down. Further, the dots dot the ‘i’s above and below the dots. By contradicting the impossibility of text reading the same upside-down, this image further illustrates the link between images we have seen and the new outline font images that are now the basis for almost all of the graphics and text that we see. The Spiral Infinity image was written in JAM (the John and Martin Language) (named for John Warnock and Martin Newell), before the PostScript PDL existed, and is copyright 1981, by Scott Kim, all rights reserved. The image is included in *Inversions* by Scott Kim, page 35, published by Key Curriculum Press, Emeryville, CA, 1996 ISBN 1-55953-280-7. Used with permission.