Sustainable IT Budgeting: A Method to Determine Not to Exceed Values for Annual Infrastructure Purchases

Steve Langer

Picture-archiving and communication systems are complex entities, but at core they consist of compute processors that are networked together to store and retrieve objects. Therein lay fundamental aspects of both performance benchmarking and predicting future costs, provided one can accurately predict trends in both exam volumes and sizes. Hence, determining the correct amount of capital to reserve annually for the information technology infrastructure can be a difficult process for the administrator of a medical center. Both exam volumes and sizes tend to increase over time. In addition, users demand more compute-intensive applications and expect exam delivery to the desktop to be ever timelier despite the increase in size. Against this, storage, compute, and networking costs tend to decrease over time for the same performance level. At the end of the day, the question of whether to budget more or less capital for next year's infrastructure is not trivial. This paper develops a methodology that uses current baseline data to predict the "ampleness" of a budget to meet future needs.

KEY WORDS: Price/performance ratio, PACS, IT budgeting

BACKGROUND

The literature is rife with papers on planning to purchase and install picture-archiving and communication systems (PACS)¹⁻⁷. However, of equal importance is the care and feeding over the life of the PACS once it is operational. There is comparatively little written on this topic. In particular, there is little to guide the financial planning for the ongoing maintenance of the PACS other than the advice of the vendor during the contract negotiations when planning for the postwarranty period.

Often, a medical center's PACS or department administrator is expected to track the annual expenses associated with running various information systems and infrastructure and predict the costs needed to maintain them in the future. This can be a challenge. To see why, consider the following simple case:

1. Annual PACS storage needs have increased 7 TB/year (where TB is terabyte) for the past 3 years.
2. Next year's storage fee should be the same as this year's—but can one be certain?
3. However, the cost/TB is dropping 25% per year. Hence, next year's storage costs should drop.
4. However, the clinic just ordered two 64-slice computed tomography (CT) scanners for next year. The average exam size is expected to grow from 300 slices to more than 2,000 per exam and exam volume is expected to increase by 7%.

Therefore, should the administrator budget more or less for storage next year?

Similar scenarios can be played out for computational resources (are new servers needed?) or networks (is it time for 1 billion bits per second to the desktop?). Is it possible to predict with

From the Mayo Clinic, Rochester, MN, USA.

Correspondence to: Steve Langer, Mayo Clinic, Rochester, MN, USA; tel: +507-266-4418; fax: +507-284-8996; e-mail: Langer.steve@mayo.edu

Copyright © 2008 by Society for Imaging Informatics in Medicine

Online publication 3 June 2008

doi: 10.1007/s10278-008-9126-z

certainly a not-to-exceed annual budget line item that will assure that resources will always keep pace with demand? With a few caveats, yes it is.

MATERIALS AND METHODS

The model for estimating infrastructure capital is based on the following logic. For the relatively simple case of storage, assume this year $\$X$ for Y TB is spent for a performance/ $\$$ ratio of one mirrored-TB/ $\$25K$. Assume next year that total data volumes (number of exams multiplied by size per exam) will grow by about 10%. Hence, storage needs for next year will be $1.1Y$ TB. Is it necessary to spend $\$1.1X$? Probably, it is not. If the storage vendor prices the product in line with market forces, the TB/ $\$$ should be increasing about 25%/year⁸⁻¹¹. Hence, spending the same dollars next year (X) should yield $1.25Y$ TB, which is more than enough to meet expanded needs.

However, what if the CTs are replaced with 64-slice models, and new protocols predict a data volume explosion of 50%? Should the budget be more than $\$X$? How much more? Doing nothing will have $\$X$ buying $1.25Y$ TB, but $1.5Y$ TB is needed. Thus, to be safe, the budget should be $\$1.3X$ for next year.

The above analysis has used the known present day adequacy of storage, plus predicted data volume growth for next year, to derive a future budget estimate—which is corrected—for the known annual growth rate of the performance/ $\$$ for that storage. This approach can be expanded to the more general case.

For any given performance requirement (storage, networking, compute power, etc), if the growth rate of the performance/ $\$$ for that item is greater than the growth rate of the requirement, then for any budget cycle beginning with sufficient funds, a constant budget allocation will always be sufficient. This is equivalent to stating that the slope of the requirement grows slower than the slope of the performance/ $\$$ graph for that resource.

Corollary: If the slope for a given requirement is greater than the performance/ $\$$ slope for that item (i.e. storage), then even if today's budget

is sufficient, at some future date, the institution will have to budget increasing dollars in succeeding years for that resource.

The model requires a few assumptions:

- A site's adjusted data volume (exam counts multiplied by exam size) can be predicted for at least the next year and preferably 2
- A site's networking requirements scale with the increasing data volume to maintain similar exam loading times for users
- A site's computational requirements scale with increasing data volumes and functionality to maintain similar user interface response for users
- Accurate data exist to quantify the growth rate of the performance/ $\$$ ratio of central processing unit (CPU), networking, storage, and other resources.

A survey of the literature can assist in building tables that can address item d above. We begin with storage. In this case, the metric of interest is megabyte/ $\$$. The following table shows what a constant dollar investment has bought in terms of storage at each time point. From this, one can plot the semi-log performance/ $\$$ curve over time and determine the annual growth rate, which is related to the slope (the greater the slope, the faster a dollar buys increased storage; Table 1).

Note that in Table 1, we have listed both the actual megabytes/ $\$$ and the $\log_{10}$ of that value. Using the year and Log values, one can plot the semi-log function of the increase in performance/ $\$$. Furthermore, using the SLOPE function in Microsoft Excel, we can determine the slope from the table of values (Microsoft, Redmond, WA, USA). Doing this provides one with two methods to determine budgets: either by comparing the slopes of the resource versus the need (from the tables) or by laying the two plots on the same graph and visually determining the relationship. In general, there are three possibilities:

- The slopes are the same (i.e., the resource and need plot are parallel), and one can just budget constant dollars over time.
- The need slope is less than the resource slope, and one can budget reduced dollars over time.
- The need slope is rising faster than the resource, and even if the need is met at present, in the future, increasing dollars will be needed.

Table 1. The Dominant PC Disk Technology Over Time, and the MB One Could Purchase for a Constant Dollar Investment (approximately \$200 in 1980 Dollars)

Year	Technology	MB/\$	Log ₁₀ (MB)
1980	MFM	0.002941176	-0.0014789
1983	MFM	0.05	1.22897
1994	EIDE	5	3.22897
2001	IDE	150	4.70609126
2005	SATA	1,000	5.53
2007	SATA	5,000	6.22897

Slope = 0.21589194. The data can be plotted on a semi-log graph, and the annual growth (slope) of MB/\$ is determined conservatively to be 0.216

Using the above information, the interested reader can look at Table 4 and predict what will be needed for storage/computer/network spending at the institution.

The next table shows the behavior for compute power. This situation is a bit different than the storage case. Depending on the installation's size, replacing an entire site's PACS workstation and server fleet annually would be a prohibitive human resource expenditure that would leave little time for anything else. Instead, computers are usually sized to initially exceed the present need, but over their life, the application demand will grow until the computers are no longer memory and CPU adequate for the tasks assigned. Hence, in the case of storage, resource consumption is continuous and needs replenishment annually; thus, funds for storage must be allocated annually. In contrast, computers tend to be ever-greened over an N year cycle (where N is typically 3–5 years). Operationally, this can be handled by encumbering $1/N$ of the total evergreen cost every year, so that on the N th year, a site has the funds to replace the computers. Hence, the challenge in this case is to

estimate the CPU performance needs over the planning cycle and purchase machines exceeding that performance at that time.

Given the preceding, what is the proper metric? As before, we seek to find a gauge whose performance evolution over time for constant dollars can be charted. A useful metric for personal computers (PCs) is MIPS/\$ (where MIPS is defined as million instructions per second). For the purpose of this table, the data comes from the history of Intel 80×86 family based PCs (Table 2)^{11–13}.

Last, we look at the information needed to build a performance/\$ plot for networking. In this area, the metric of interest is Mbps/\$ where Mbps is million bits per second (Table 3)^{11,14,15}.

RESULTS

It is instructive to look at the isolated case of CT at our institution. The following table shows the disruptive effect of acquiring more than ten 64-slice CT scanners in 1 year (2004). Note that from 2002 to 2004, slice counts increased approximate-

Table 2. The Changing Performance/\$ for Constant Dollars (Approximately \$2000 In 1980) For PCs Based on the Intel Chip Family from 1980 to 2004

Year	Device	MIPS/\$
1980	2 MHz 8088	0.008
1986	8 MHz 286	0.32
1991	33 MHz 386	5.94
1994	66 MHz 486	33
1996	90 MHz Pentium	71.1
1999	400 MHz Pentium II	400
2004	2000 MHz Xeon	4,000

A fit shows an annual MIPS/\$ increase of 0.238

Table 3. The Networking Speed Over Time (In 1980 Dollars) for a Constant Price Per Network Adaptor

Year	LAN technology	Mbps/\$
1980	Thicknet 3	0.03030303
1983	Thinnet 10	0.15384615
1987	10 Base2 shared	0.25
1990	10 BaseT switched	0.45454545
1995	100 BaseT switched	5
1999	1,000BaseT	50
2003	10,000 BaseT	40.6504065

The growth rate for Mbps/\$ is 0.145.

ly 4 million slices per year in concert with normal exam count growth, but from 2004 to 2005, slice counts surged ahead 9 million images per year. Since 2005, incremental slice counts have been exploding 20 million images per year. This is not from mere exam count increases but exam protocols surging from an average of 300 slices to more than 1,000 slices per exam, with exceptions up to 6,000 slices not uncommon (Table 4).

To put all data in context, it is helpful to plot all the above tables on one graph. This can be done by:

- Plotting years against the $\log_{10}$ of the performance/\$ metric for Tables 1, 2, and 3
- Adjusting the y -intercepts for Tables 1, 2, and 3 to have their origin at zero at a common starting date
- Adding the semi-log plot of CT growth to the above plot

This permits a single plot to indicate all the trends for infrastructure spending for PACS, as compared to the annual growth rate for CT. It is evident that a networking challenge in the near future will require increased funding; one can see this both visually and by noting the slopes in the captions (CT growth is increasing at 0.193, while networking is increasing at only 0.145). Assuming that users will not tolerate a linear increase in loading time with a linear increase in exam size,

Table 4. The Annual Production of CT Images at our Institution

Year	CT size	$\log_{10}$ (CT size)
1997	2,760,759	6.441028497
1998	3,485,113	6.542216864
1999	4,520,836	6.655218753
2000	5,544,986	6.743900454
2001	7,307,608	6.863775243
2002	9,746,599	6.988853098
2003	13,011,629	7.114331672
2004	17,565,699	7.244665437
2005	26,368,808	7.421090498
2006	46,516,823	7.667610046
2007	63,911,019	7.805575742

Slope = 0.135, slope since 2004 = 0.192. One can compute the growth rate over the period and finds that since 2004 it is 0.193

Fig 1. In this figure, all resource plots have been normalized to an origin of 1995 (year 15), with their y -offsets adjusted to the performance/\$ origin. This permits their slopes to be directly visually compared with each other and with the onset of our CT volume records. We then superimpose the CT growth rate on the plot, beginning with data from 1997. Of particular interest is the change in slope at about the year 2004. This corresponds to the conversion of half of our site's CT fleet to 64 slice scanners. It shows that the rate of growth is itself growing!

we are directed to increase network bandwidth. It is also possible that PACS workstations at some institutions will need more random access memory to support the larger data sets. Fortunately, our institution foresaw this trend in 2005 since our PACS viewer is a fat client and sized the PACS PCs to 64-bit hardware using the XP-64 operating system (XP64, Microsoft; Fig. 1).

DISCUSSION AND CONCLUSIONS

As mentioned earlier, the data for resource growth have to be accurate for one's own institution. Further, it may also be necessary for a site to derive its own performance/\$ metrics if their suppliers do not pass on the savings of the broader information technology (IT) markets. We were fortunate to negotiate in the PACS purchasing contract a stipulation that we would acquire hardware directly and eliminate the "middleman." All the forgoing is just another way of saying, "Please just don't copy our results, your mileage will vary." With these caveats in mind, however, we have found the described method a useful model for predicting future IT spending.

ACKNOWLEDGMENTS

The author would like to thank the following for useful data: Jason Tjelta and Ken Persons, both from the author's home institution's IT department.

REFERENCES

1. Siegel EL, Reiner BI, Siddiqui KM: Ten filmless years and ten lessons: a 10th-anniversary retrospective from the Baltimore VA Medical Center. *J Am Coll Radiol* 1(11):824–833, 2004
2. Samei E, Seibert JA, Andriole K, Badano A, Crawford J, Reiner B, Flynn MJ, Chang P: AAPM/RSNA tutorial on equipment selection: PACS equipment overview: general guidelines for purchasing and acceptance testing of PACS equipment. *Radiographics* 24(1):313–334, 2004
3. Reiner B, Siegel E: Understanding financing options for PACS implementation. Picture archiving and communication systems. *J Digit Imaging* 13(2):49–54, 2000
4. Stewart BK: PACS: a phased implementation strategy. *Adm Radiol* 14(11):10–16, 1995
5. Dwyer SJ, Templeton AW, Martin NL, Cook LT, Lee KR, Levine E, Batnitzky S, Preston DF, Rosenthal SJ, Price HI, Anderson WH, Tarlton MA, Faszold S: Cost of managing digital diagnostic images for a 614-bed hospital. *J Digit Imaging* 15(4):255–60, 2002
6. Langer SG, Wang J: A goal based cost–benefit analysis for film vs. filmless radiology departments. *J Digit Imaging* 9(3):104–112, 1996
7. Langer SG, Wang J: Purchasing guidelines for information system software in a filmless radiology environment. *J Digit Imaging* 9(3):113–118, 1996
8. Daney C: Digital storage technology. In: Science and Reason. 2007. Available at: <http://scienceandreason.blogspot.com/2007/10/digital-storage-technology.html> (last accessed February 2008)
9. Kurzweil R: The law of accelerating returns. In: Kurzweil AI. 2001. Available at: <http://www.kurzweilai.net/articles/art0134.html?printable=1> (last accessed February 2008)
10. Little Tech Shoppe: Historical Notes About the Cost of Hard Drive Storage Space. 1981. Available at: <http://www.littletechshoppe.com/ns1625/winchest.html> (last accessed February 2008)
11. Panzer-Steindel B: CPU, Disk and Memory Issues. CERN Large Hadron Collider Grid Computing Project. 2005. Available at: http://lcg-computing-fabric.web.cern.ch/LCG-Computing-Fabric/fabric_presentations/CPU_disk_and_memory.doc (last accessed February 2008)
12. Moore's law. In: Wikipedia. Available at http://en.wikipedia.org/wiki/Moore's_law (last accessed February 2008).
13. The Linux Documentation Project: The frequently asked questions about BogoMips. In: BogoMips Mini-Howto. Available at <http://tldp.org/HOWTO/BogoMips/bogo-faq.html> (last accessed February 2008)
14. IEEE 802.3. In: Wikipedia. Available at: http://en.wikipedia.org/wiki/IEEE_802.3 (last accessed February 2008)
15. Dealttime. Available at: http://www0.dealttime.com/xPP-network_adapters (last accessed February 2008)