The Technology Economics of the Mainframe, Part 3: New Metrics and Insights for a Mobile World

Dr. Howard A. Rubin
CEO and Founder, Rubin Worldwide
Professor Emeritus City University of New York
MIT CISR Associate
Gartner Senior Advisor
howard.rubin@rubinworldwide.com
www.rubinworldwide.com

Abstract

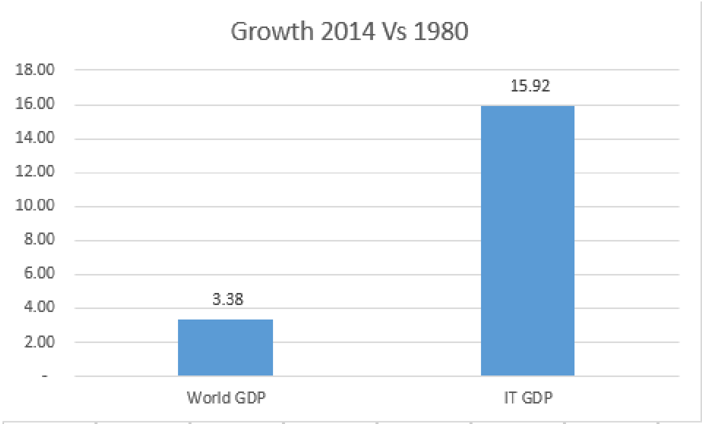
The global economy, whether considered from a business or personal perspective, is a technology economy; its foundation is computational, storage and network capacity and capabilities that are intensifying day after day. Spending for this technology has grown at a rate of more than 4 times that of the world's GDP over the past 35 years. Looking ahead, technology intensity fueled by an explosion in mobile transactions is at a tipping point with an imminent need for a massive increase in processing power. Therefore, selection of computing platforms is now more critical than ever, for reasons of economics as much as for traditional IT infrastructure criteria. An ongoing body of research begun in 2010 has shed light on the economic implications of platform choices and their impact on business. Over the course of this research, growth in mobile fueled workloads has driven requirements for response times, scalability, security, reliability, processing agility and economic efficiency to new levels. This paper explores the forces driving this technology intensity on the basis of new metrics that provide fresh insights, with a specific focus on the economics of various platforms and architectures.

About Rubin Worldwide

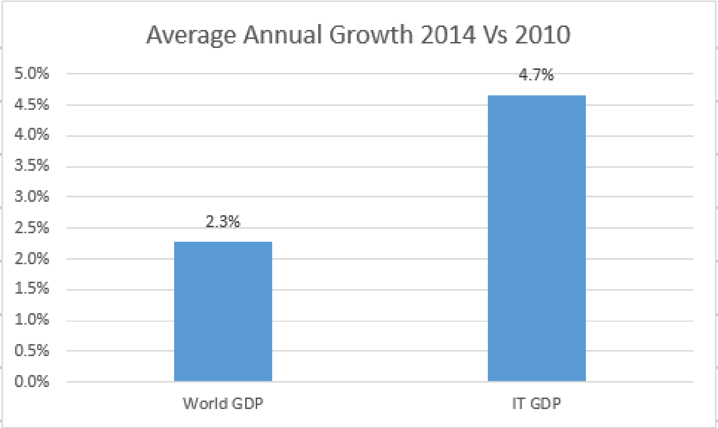
Rubin Worldwide (www.rubinworldwide.com) is a boutique research and consultancy firm with a unique focus on "technology economics". Since 1982, the founder and CEO Dr. Howard Rubin has been collecting data on the interplay and impact of technology on the global competitiveness of nations, the dynamics of technology investment and its management within companies (and government), and technology costs themselves. In collaboration with Jed Rubin in the mid-1990's, Dr. Rubin started to produce reports on global trends and patterns which now are considered the foundation of the field of technology economics. Rubin Worldwide's data and analytics are used by many of the world's governments, premiere companies, analyst firms, and consultancies to explore the evolving world of technology economics and its impact as a competitive weapon.

Technology Intensity Is Increasing in Our Technology Economy

There is no denying that we live in a world of increasing *technology intensity*: the amount of computational capacity (processor cycles, storage, network bandwidth, etc.) needed to support business growth, global economic activity and indeed our personal lives. Here is the evidence: During the period from 1980 to 2014 while the world's GDP has grown about 3.4 times, from $22T to $74T, the technology expense to support that GDP grew almost 16 times, from $335B to more than $5T.



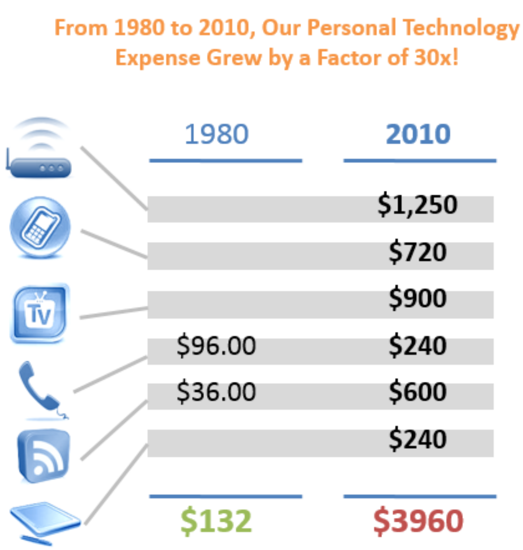
In terms of growth rates, over this period as the World GDP grew at a rate of 2.3% per year, the "IT GDP" has grown at an average rate of 4.7%.



Because consumer behavior is a major economic driver, it is important to also look at what this means from the perspective of the individual. Between 1980 and 2014, global technology spending per person (or IT GDP per Person) has grown nearly tenfold, from $76 to $726.

	World GDP ($B)	IT GDP ($B)	World Population (B)	World GDP$ per Person	IT GDP$ per Person
1980	$ 22,195	$ 333	4.4	$ 5,044	$ 76
2014	$ 75,000	$ 5,300	7.3	$ 10,274	$ 726
Ratio of 2014 to 1980	3.38	15.92	1.66	2.04	9.60

From a U.S. perspective, during the period when U.S. average personal income rose from about $17,000 to $46,000, personal technology expense increased from $132 annually to almost $4,000 – from less than one percent (0.78%) of income to more than eight percent (8.6%)!

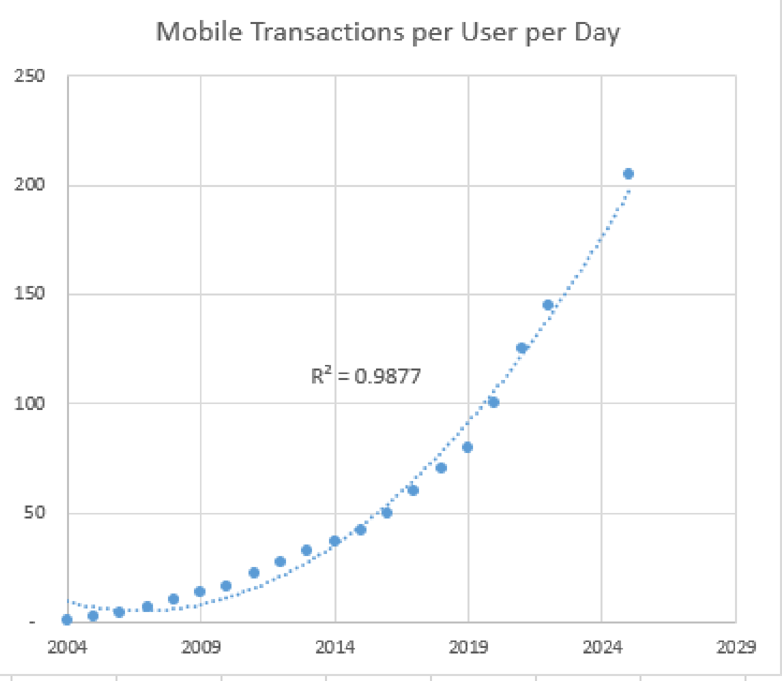


Mobile Is the New Driver of Technology Intensity

Now consider the role of mobile computing in technology intensity. Here are some facts:

- As mobile adoption grows, consumers are driving an exponential increase in mobile transactions (a "transaction" being a user-generated request for processing that results in multiple system interactions).
- In 2004, there was less than 1 mobile transaction per day per mobile user.
- Ten years later, in 2014, that number has grown to at least 37 transactions per day per mobile user.

Looking ahead to 2025, it is easy to foresee a future state approaching 5X that level… in excess of 200 mobile transactions per day per user – and with 2 billion mobile users that is a total of 400 billion user transactions per day globally.



What are the implications of this explosive mobile growth on technology intensity? What level of computational capacity will be needed to support the global economy, businesses, government and our personal lives? What will be the cost of this capacity? And even more important, what is the economic impact?

Many consider Moore's Law to be the ultimate control mechanism that keeps costs reasonable as demand for information technology increases. Countering Moore's Law, however, is Jevons Paradox, which says technology improvements that improve the efficiency of a resource actually increase demand at a rate faster than productivity. So it is possible that Jevons explains the phenomena we are seeing in global bandwidth, computing power and mobile computing: demand will outstrip Moore's Law's ability to reduce cost. This is most evident in the growth and back-end impact of mobile computing.

First, consider that each mobile transaction triggers a cascade of related events across systems. These events can include such things as comparison to past purchases, encryption, bank to bank reconciliation, customer loyalty and rewards program updates, and many other examples. This phenomenon is known as a "starburst effect," where a single transaction triggers as many as 100 additional system interactions. Consequently those 37 transactions cited earlier may produce as many as 3,700 interactions per day – all initiated by a single user.

Next, consider that by 2018, each mobile consumer is expected to drive 5,000 or more system interactions per day. By 2025 with projected growth in mobile users reaching 2 billion, this could to grow to an astounding 20,000 systems interactions per day per user, resulting in 40 trillion events per day!

Moore's Law cannot even come close to driving costs down (and power up) in the face of this demand increase.

Finally, the rise of technology intensity to date, the exponential impact of mobile and the unlikelihood that Moore's Law will "save us" make it even more critical that we understand the implications of computing platform choices being made right now. Those platform choices really boil down to 1) mainframe, 2) commodity servers and, perhaps, 3) the "cloud."

Platform Economics: A Tale of Two Banks

Let's start with a real example of two banks of similar size (2009 revenue of about $20B) and virtually identical business mixes and geographic footprints. The core processing platform deployed by the two banks in 2009 indicated very different profiles and philosophies – Bank A was mainframe "heavy" and Bank B was commodity server "heavy":

	MIPS	Servers
Bank A	60,000	22,500
Bank B	30,000	45,000

For Bank A and Bank B during the period 2009 to 2014, revenue stayed flat but the need for computation resources doubled as the wind shear forces of regulation, security, compliance, mobile, and big data analytics drove up technology intensity. By 2014 their profiles had changed:

	MIPS	Servers
Bank A	120,000	22,500
Bank B	30,000	90,000

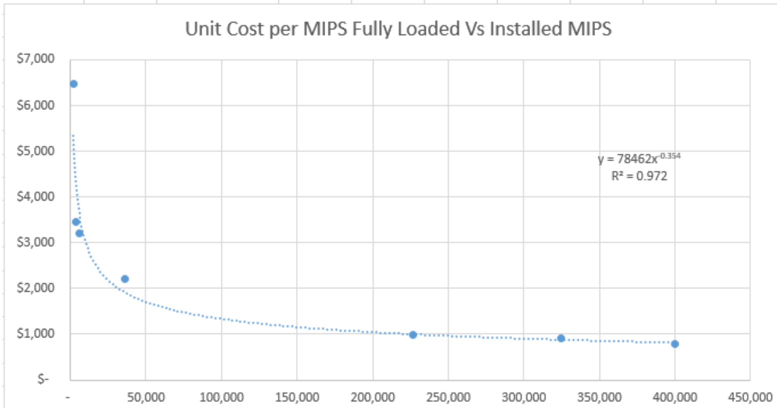
Essentially, one held the line of mainframe growth and the other commodity server growth.

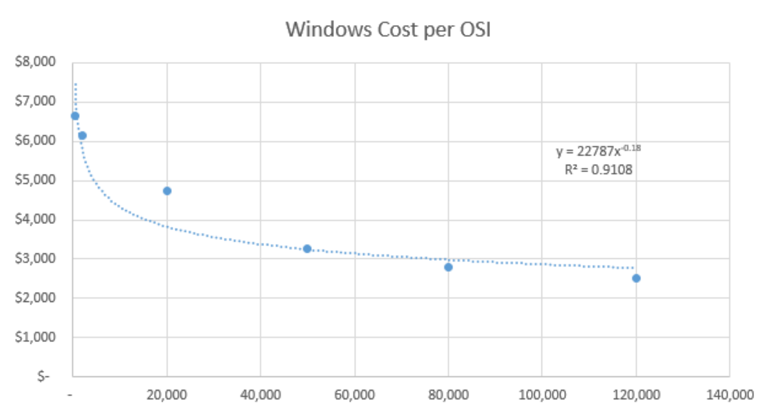
While this difference is apparent in the counts of MIPS and servers, the economic profile provides deeper insights into some of the basics of platform economics.

		2009				**2014**		
		Count	**Unit Cost**	**Total Cost**		**Count**	**Unit Cost**	**Total Cost**
Bank A	**MIPS**	60,000	$ 1,597	$ 95,795,494		120,000	$ 1,249	$ 149,903,046
	Servers	22,500	$ 3,448	$ 77,577,988		22,500	$ 3,448	$ 77,577,988
			MIPS + Servers	$ 173,373,482			**MIPS + Servers**	$ 227,481,035
			Total Infra	$ 780,180,670			**Total Infra**	$ 1,023,664,656
Bank B	**MIPS**	30,000	$ 2,041	$ 61,218,080		30,000	$ 2,041	$ 61,218,080
	Servers	45,000	$ 2,973	$ 133,766,723		90,000	$ 2,563	$ 230,652,234
			MIPS + Servers	$ 194,984,803			**MIPS + Servers**	$ 291,870,314
			Total Infra	$ 877,431,616			**Total Infra**	$ 1,313,416,415

The doubling of computing power between 2009 and 2014 in the mainframe heavy organization (Bank A) resulted in a 30% increase in core infrastructure costs. In Bank B with its commodity server heavy profile, this doubling resulted in 49% more cost.

The underlying reason for this is inherent in the power curves of unit costs and scale for mainframe and commodity server environments. Within the ranges experienced by most businesses today, the mainframe is inherently more scalable because decreases in unit cost offset volumetric changes more so than for commodity servers.





While the aforementioned data provides a technical and IT cost view of the mainframe heavy vs. commodity server heavy phenomena, a business oriented view is more revealing. Consider a new metric: Dollars of Income Supported per Dollar of Infrastructure Expense. The data for our two banks (normalized to the same level of income) is shown below with Bank A exhibiting higher Income Support Infrastructure Economic Efficiency than Bank B.

	$ Income per $ Infrastructure Cost		
	2009	**2014**	**% Change**
Bank A	$ 13.84	$ 10.55	76%
Bank B	$ 12.31	$ 8.22	67%
Average Bank	$ 5.44	$ 3.86	71%

Bank A's (mainframe heavy) ratio of business income to infrastructure costs was higher than Bank B (commodity server heavy) in 2009 and remained so in 2014. In fact, Bank B's commodity server heavy infrastructure resulted in a 33% drop in the ratio of business income to infrastructure cost in 2014, while for Bank A the change was 24%. In fact, Bank B fared worse than the average bank, which exhibited a drop of 29%.

New Insights from a New Metric

This real-life example indicates that the cost structure of mainframe computing is favorable and becomes more so as technology intensity increases. Initial insights from his new metric indicate that mainframe computing may offer higher infrastructure economic efficiency when viewed in business terms, e.g. Income per Infrastructure $. Many IT organizations fail to see this because they simply look at the mainframe in terms of its overall cost pool in contrast to the typical view of commodity servers, which is on a per-unit basis; this provides a badly distorted view.

A complete understanding of the mainframe's computational and economic relevance in the context of its contribution to business performance is critical. This is shown clearly by examining the metric of Income Supported per Dollar of Infrastructure Expense in areas other than Banking and Financial Services.

In all but 2 of 19 sectors studied, mainframe heavy organizations supported more income per dollar of infrastructure expense than peers that were commodity server heavy.

	Income Supported per $ of Infrastructure Expense	
	Mainframe Heavy	Commodity Server Heavy
Banking & Financial Services	$5.54	$3.88
Chemicals	$13.34	$10.67
Construction, Materials, and Natural Resources	$40.63	$30.47
Consumer Products	$8.00	$5.76
Education	$2.38	$2.26
Energy	$46.53	$36.29
Food & Beverage Processing	$29.56	$23.35
Healthcare Providers	$2.33	$1.75
Industrial Electronics and Electrical Equipment	$13.24	$9.54
Industrial Manufacturing	$5.92	$4.14
Insurance	$1.63	$1.24
Media and Entertainment	$9.97	$9.07
Pharmaceuticals, Life Sciences, and Medical Products	$11.13	$7.90
Professional Services	$6.62	$4.64
Retail and Wholesale	$22.77	$17.76
Software Publishing and Internet Services	$0.27	$0.33
Telecommunications	$5.71	$4.00
Transportation	$13.15	$10.26
Utilities	$13.97	$11.45
Full Database	$13.35	$10.55

More Business Insights from “Old” Metrics

Past studies (2010 and 2012) of platform economics support the current finding of increased economic efficiency for mainframe heavy infrastructures where mainframe usage is appropriate. An updated analysis of IT Cost of Goods across 15 sectors supports this hypothesis and indicates that further leverage has been attained in the period 2012 to 2014.

2012 - 2014 Analysis							
Industry	Measure	Average IT Cost of Goods	Mainframe Heavy	Commodity Server Heavy	% Mainframe Cost Less than Server	2010-2011 Differential	Change
Bank	Per Teller Transaction	$ 0.030	$ 0.125	$ 0.401	69%	67%	2%
Mortgage	Per Approved Loan	$ 295.300	$ 100.200	$ 358.400	72%	68%	4%
Credit Card	Per Transaction	$ 0.138	$ 0.094	$ 0.192	51%	48%	3%
Railroads	Per Ton Mile	$ 0.0011	$ 0.0012	$ 0.002	39%	36%	2%
Armed Service	Per Person	$ 9,410.000	$ 7,124.000	$ 12,544.000	43%	35%	9%
Automotive	Per Vehicle	$ 382.000	$ 279.000	$ 413.000	32%	31%	1%
Retail	Per Store/Door	$ 560,266.000	$ 453,444.000	$ 675,899.000	33%	27%	6%
Utilities	Per MegaWatt Hour	$ 2.580	$ 2,499.000	$ 3,345.000	25%	19%	6%
Hospitals	Per Bed per Day	$ 82.880	$ 62.320	$ 91.560	32%	27%	5%
Oil & Gas	Per Barrel of Oil	$ 2.330	$ 1.800	$ 2.610	31%	28%	3%
Consulting	Per Consultant	$ 58,650.000	$ 48,766.000	$ 68,100.000	28%	28%	1%
Trucking	Per Road Mile	$ 0.185	$ 0.160	$ 0.225	29%	20%	9%
Airlines	Per Passenger Mile	$ 0.009	$ 0.007	$ 0.010	36%	30%	6%
Chemicals	Per Patent	$ 66,588.000	$ 58,922.000	$ 68,566.000	14%	10%	4%
Web Sites	Per Search	$ 0.040	$ 0.042	$ 0.038	-11%	-8%	-2%
				Average	35%	31%	4%

The gap between mainframe and commodity server heavy organizations increased another 4% on a cost of goods basis due to the ability of the mainframe environment to scale more efficiently than that of commidty servers. And, in more distinct business terms, the differential in IT costs on a per-transaction basis is substantial.

A New Perspective: Climate Change and Platform Choices

The digital economy in total consumes more than 10% of the world’s electricity. For instance, an iPhone with its connectivity, recharging, and data storage consumes as much energy as a medium sized refrigerator (Time Magazine, 8/14/2013). This makes for an interesting perspective on platform choices and economics from a carbon footprint view.

Conventional wisdom would indicate that a mainframe consumes more power than a commodity server. However, translating processing power into a transaction/business view provides what might be considered a climate change impact perspective. See the examples below as an illustration.

	LBS of CO2	
	Mainframe Heavy	**Commodity Server Heavy**
Credit Card	0.04	0.23
Loan Organization	0.14	0.38
Process and Maintain $1M Loar	103.00	709.00

Because of its ability to process high transaction volumes at extremely high economic efficiency, the mainframe is likely to have a relatively smaller carbon footprint again due to its scalability – its inherent ability to add computational cycles without a linear increase in power consumption more typical of commodity servers.

Clearly, much more needs to be done to explore this aspect of platform economics, but in the future this may even be a product differentiator for the consumer marketplace.

Synthesis: The "Perfect Storm" for Mainframe Economics

Combining the results of prior research with the new analyses contained in this paper, it is clear that we are entering a new technology era. A global thirst for computing capacity threatens to outstrip our ability to supply it at an acceptable price. The vagaries of the global economy, increasing digital security concerns and the mobile revolution add further to this "perfect storm" of technology intensity.

Our "new math" of technology economics indicates that the leverage and competitive advantage an organization can gain from a mainframe heavy infrastructure can be critically important to its success. Considering that an IBM z13 can perform 2.8 billion transactions per day – the equivalent of 100 retail Cyber Mondays – in addition to its demonstrated economic efficiency in business terms, its inherent high levels of security and scalability, and even its potential for "greenness," the mainframe is a platform well suited to the emerging mobile driven world.