Transformation of Business and Financial Platforms: What We Must Understand to Renew Existing Systems

Kip M. Twitchell
IBM Consulting
IBM
Phoenix, Arizona
Kip.Twitchell@us.ibm.com

***Abstract*— Transforming existing business and financial applications is very challenging, particularly at scale where the stakes are highest. The data in them has been gathered and certified by auditors at great cost. The systems developed and tuned typically over years if not decades. A wrong turn can threaten the market value and even the existence of the organization. True ledgers not only capture transactions, but also maintain balances or positions over time.**

For blockchain to fulfill its potential to transform the world's business systems, most of which were automated before the Web was invented, developers must understand:

- **That while financial reporting on business events is the bedrock of modern capital markets,**
- **Reporting purely on business events is inefficient; efficient systems include posting to balances**
- **Example data structures required can be found in analogies to the daybook, the journal, and ledger of historical bookkeeping**

Keywords—blockchain, ledgers, accounting, bookkeeping, scale, daybook, journal entries, legacy systems, financial data

I. Introduction

Blockchain's potential of reinventing the world's financial and business system processes has had mixed results to date. Recent trends focusing on reduced costs, greater scale, and working within existing government and regulatory frameworks portent greater momentum in these efforts.

With a professional focus for over three decades on reinventing these systems, and a student of blockchain for over half a decade, it is clear actual reinvention requires greater understanding of the principles behind today's existing financial and business systems.

II. Financial Information: The Bedrock of Capital Markets

Modern financial markets and banking are fundamentally dependent upon financial reporting. Release of financial information moves trillions of units of currency every business day. The foundation for these systems has developed over centuries. Crosby argues the modern world developed based upon quantification and measurement, leading to incremental improvements in many domains.

"Double-entry bookkeeping was and is a means of soaking up and holding in suspension and then arranging and making sense out of masses of data that previously had been spilled and lost…. Today computers compute faster than friar Pacioli would ever have dreamed possible, but they do so within the same framework (accounts payable, accounts receivable, and all) as he did. The efficient friar taught us how to oblige grocery stores and nations, which are always whizzing about like hyperactive children, to stand still and be measured….

"In the past seven centuries bookkeeping has done more to shape the perceptions of more bright minds than any single innovation in philosophy or science…

"In practical terms, the new approach was simply this: reduce what you are thinking about to the minimum required by its definition; visualize it on paper, or at least in your mind, be it the fluctuation of wool prices at the Champagne fairs or the course of Mars through the heavens, and divide it, either in fact or imagination, into equal quanta. Then you can measure it, that is, count the quanta." [1]

Four decades ago, McCarthy noted that the underlying data in financial affairs is primarily economic or business events.

"*Economic events* are defined by Yu [1976, p. 256] as "a class of phenomena which reflect changes in scarce means [economic resources] resulting from production, exchange, consumption, and distribution." For both Ijiri [19751 and Sorter [1969], economic events constitute the critical information elements of an accounting system…. When the [REA] model…is considered in its maximum form of temporal generality, (without considering implementation cost), event descriptions would be maintained perpetually as base elements of the conceptual schema. That is, detailed descriptions of all transactions would be stored indefinitely in disaggregated, individual form." [2]

The fundamental data structure at the bottom of most financial reporting is a financial transaction. The data that moves the markets daily is primarily an accumulation of business events. Accumulation is the process of measurement Crosby describes.

III. Balance or Positions and Computer Efficiencies

McCarthy posited a radical redefinition of financial system architectures:

"It is a primary contention of this paper that the semantic modeling of accounting object systems should not include elements of double-entry bookkeeping such as debits, credits, and accounts. As noted previously by both Everest and Weber [1977] and McCarthy [1979], these elements are artifacts associated with journals and ledgers (that is, they are simply mechanisms for manually storing and transmitting data). As such, they are not essential aspects of an accounting system. It is possible to capture the essence of what accountants do and what things they account for by modeling economic phenomena directly in the conceptual schema. Any double-entry manipulations desired by particular users can then be effected only in the external schemata presented to those users." [3]

In other words, with unlimited computing capacity and the ability to manage the complexity, we could, theoretically, accumulate all the transactional data to produce all the financial reports at any point in time.

Said another way, the balance sheet of IBM or any other entity at any time includes the net effect of every financial transaction since the first, such as an initial capital infusion. With all transactions stored in an electronic format, computer programs could be written to repeat the steps taken to produce any period's financial records. The data could also be used to produce unlimited numbers of other perspectives, not precluded by tremendous aggregation and the imposition of accounting data structures and processes of past periods.

Yet McCarthy's paper also recognized how radical this approach was when attempting to implement the system practically at scale.

"In practice, [detailed descriptions of all transactions…stored indefinitely in disaggregated, individual form] is an unrealistic assumption…. " [4]

"During design of an actual system, quite obviously, consideration of both decision usefulness and storage costs would temper these requirements somewhat and identify places where temporal summation of event data might be appropriate..." [5]

Effectively, while business events qualify as the quanta in Crosby's bookkeeping. one can ask how to do that "count the quanta" in the most efficient way? This question deserves much more rigorous thinking to enable blockchain's ambition. Blockchain's storage of only the business event data reflects McCarthy's "unrealistic assumption" and is an impediment to scale, cost-effectiveness, and usefulness.

Existing systems continue to function as they do in some cases because of the efficiencies provided by true ledgers, which "post" transactions to balances to maintain a running position. Balances give a point in time answer to any question. Almost all analytical processes begin by analyzing a balance. Balances are also efficient for computer processes. They eliminate the need to deal with historical transactions to state current positions.

Consider the issues as described by me in 2018:

"This chart [Fig. 1] shows the well-known trends of increasing computing spending and declining computing costs

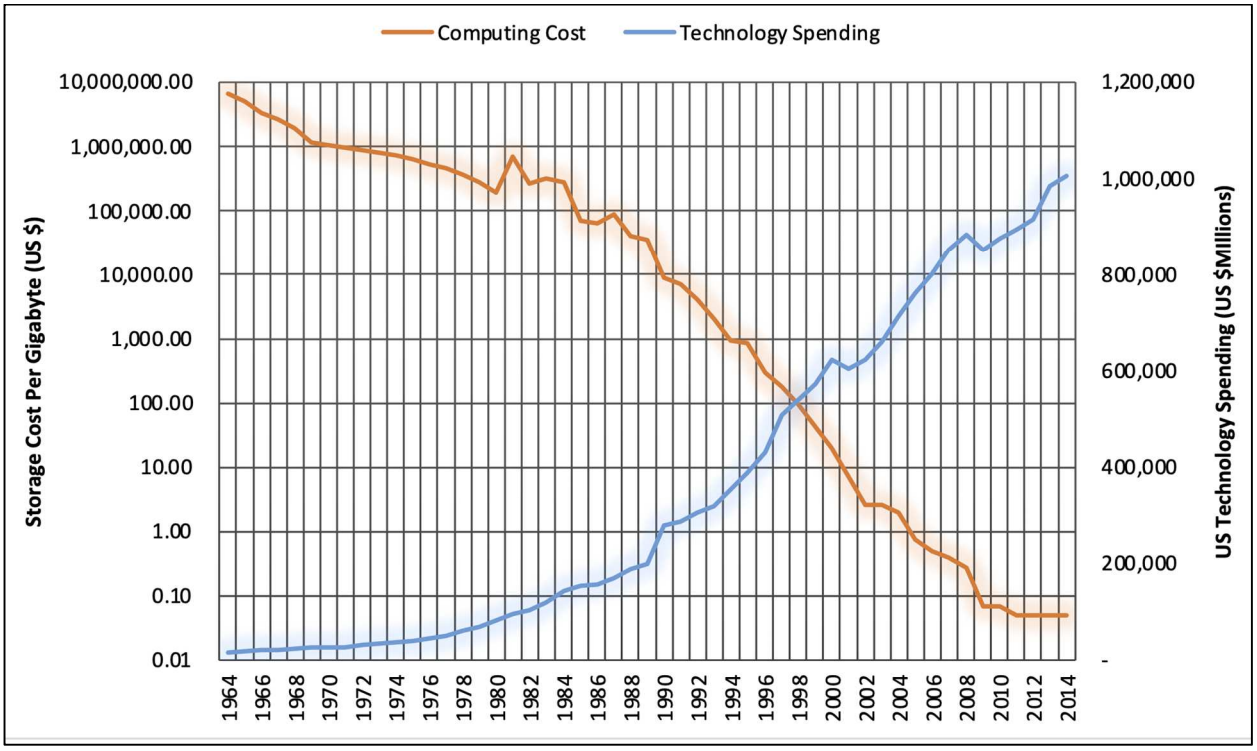

Figure 1. Compute cost vs. technology spending over 50 years [6] [7]

.

from 1964 to 2014. In today's world of ubiquitous computing capacity, we might productively consider how it is possible to organize data ...in [the limited memory of a] 1964 10-million-dollar computer....

"The answer is to use ledgers, with the associated balances. Computing resources are highly sensitive to transaction volumes in processing. Balances eliminate the need to reprocess historical transactions to get to current positions. Each posting cycle is limited to the transactions since the last update plus the current balances, producing a new set of balances reflecting the updated position." [8]

The following quote about psychology can be applied to the trade-offs between transactions and balances:

"In the practical use of our intellect, forgetting is as important a function as recollecting…. If we remembered everything, we should on most occasions be as ill off as if we remembered nothing. It would take as long for us to recall a space of time as it took the original time to elapse, and we should never get ahead with our thinking….

"…We thus reach the paradoxical result that one condition of remembering is that we should forget. Without totally forgetting a prodigious number of states of consciousness, and momentarily forgetting a large number, we could not remember at all…. " [9]

As efficient as summaries, balances or positions are, they are secondary evidence. If they are not supported by transactional detail, balances are irrelevant. It's also important to recognize, balances have downsides:

- A balance is effectively a duplication of selected attributes and accumulated amounts of the transaction data. As such it often requires reconciliation to validate that it is correct.
- Balances typically do not contain all transactional detail, and thus cannot answer all information needs transactions can.
- Creating or updating balances requires some level of processing beyond transaction capture

IV. The Data Structures of Bookkeeping

A data supply chain produces the positions required to measure the business world. It turns disconnected business events into measurable, actionable information. Data supply chains that use bookkeeping principles can produce consistent, comparable, and understandable results over long periods of time. They can involve very sophisticated information models, like accounting. [10]

The data structures needed for efficient information systems can be understood by analyzing historical bookkeeping practices.

Bookkeeping was first documented in vernacular Italian to be accessible to business people rather than the academic Latin by Luca Pacioli in his 1494 work, *Summary of Arithmetic*. [11] A small section of the book describes three books to be kept in bookkeeping, The Day Book, the Journal, and the Ledger. All three of these "books" typically exist in modern financial systems.

A. The Daybook

The daybook is the original transaction repository. It was often written more as text than structured data, recording dates, names, products, amounts, and any other attributes of the sale.

These records still exist in today's business systems and are known as the source system, or original book of record. They contain the narrative contracts and sales terms along with transactions records. Copies of daybook entries are exchanged with counterparties as receipts.

Attributes of the daybook include:

- Chronological organization and format
- Often unstructured
- Business-event (transaction) oriented, focused on being comprehensive

Blockchain is well suited and often used to store of this type of data, either directly or in separate structures the hash of which can be verified by the blockchain.

B. The Journal

Historically, the bookkeeping clerk would periodically use the daybook to create entries in the Journal, creating "journal entries." The process consisted of turning unstructured or semi-structured data into structured data of values chosen for later reporting purposes, such as the customer or the account.

Today, many choices are available when constructing journals. For example, bookkeeping can include very simple booking models that do not include what McCarthy calls "duality" or a debit/credit representation. The choice made when building a journal is based upon the reporting requirements of the data..

In modern systems, this process is typically either part of the transaction system that interfaces to the accounting system or is part of the accounting system. It has come to be called Accounting Rules Engines.

Attributes of the journal include:

- Chronological organization and format
- Structured
- Business-event (journal entry) oriented, informed by reporting and analytical needs

C. The Ledger

Periodically, bookkeepers take the journal and "post" the journal entries to the ledger.

The ledger, rather than being organized chronologically, is organized by the data values in the journals. For example, a ledger page was often dedicated to each customer, the pages ordered by customer name or account number. Each customer's page contained the carry forward balance from the prior ledger page. Journal entries related to that customer are copied below the carry-forward balance. At the end of the posting process, a

new balance representing the prior balance plus subsequent journal entries is calculated.

Ledgers continue to exist for most businesses with significant size. Typically, these organizations have more than one ledger for customer, vendor, employee or other contract balances. Enterprise perspectives are maintained in the general ledger, a system almost universally required for all businesses.

Attributes of the ledger include:

- Non-chronological; organized by reporting data elements
- Structured
- Position-, or balance-oriented, providing timely and efficient answers to information needs

Blockchain is not well suited for this data structure. It has nearly no provision for the non-chronological requirement. Without these types of ledger abilities, blockchain will not scale and so cannot replace many of the world's business systems efficiently.

An approach to balancing the efficiencies of balances with the accuracy and flexibility of transactions is presented in Twitchell, K. "*A Proposed Approach to Minimum Costs for Financial System Data Maintenance.*"

V. Conclusion

Financial and business systems were some of the first systems automated. The cost benefit of automation was clear: buildings full of people with adding machines and paper ledgers were replaced by a single machine which almost never made mathematical mistakes.

Financial data in some of the largest systems has been consistently maintained since first automated. Control and audit tests over the years, at times costing millions of dollars or their equivalent, provide assurance that the data is appropriately captured and managed. Inaccurate data can threaten the existence of the organization through bankruptcy and other legal consequences.

These systems have been tuned over decades as well, and costs tracked in some ways at a detail level, with small incremental savings consistently captured and accumulated. Yet the rate of savings is continuing to decline, and the costs are perhaps approaching the minimum possible, particularly given spend required to achieve more.

Innovation in many of these systems has been limited over the last twenty or more years as effort has been focused on other IT systems, including the internet, mobile, social, AI, and others. But replacing existing systems, which were funded by huge automation savings, is much more challenging than digitizing new things.

The energy provided by the prospect of change in these systems by blockchain is the first step in innovation. But lack of understanding—zeal without knowledge—will not accomplish the objectives.

Acknowledgment

Appreciation expressed to Moayad Aloqaily and Yaser Jararweh and for the invitation to present and publish this paper, and to Carlton Doe, James Hladyshewsky, Torje Vingen Sunde, and James Belding for providing feedback on it.

References

[1] Alfred W. Crosby. *The Measure of Reality, Quantification and Western Society, 1250-1600.* xiv + 245 pp., Illus., index. Cambridge/New York: Cambridge University Press, 1997. Pages 200, 220, 221, 228.

[2] William E. McCarthy, "The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment" *The Accounting Review*, Vol. 57, No. 3 (Jul. 1982), pp. 554-578. Published By: American Accounting Association. p. 562

[3] Ibid. pp. 559-560

[4] Ibid., p. 562

[5] Ibid., p. 572

[6] GDP Technology Spending has been used to measure increased computing resources. With assistance from Edward T. Morgan, Chief, Industry Sector Division, Bureau of Economic Analysis, U.S. Department of Commerce, from US Federal Bureau of Economic Analysis GDP "Value Added by Industry" Release Date: April 19, 2018, specifically the Addenda Row for the years 1997 – 2017 "Information-communications-technology-producing industries: Consists of computer and electronic product manufacturing (excluding navigational, measuring, electromedical, and control instruments manufacturing); software publishers; broadcasting and telecommunications; data processing, hosting and related services; internet publishing and broadcasting and web search portals; and computer systems design and related services." Accessed May 22, 2018. The Technology Addenda data is not calculated before 1997; to estimate data prior to 1997 the chart includes the same combination of rows used to calculate the addenda data post 1997, based upon row titles value and addenda description, from the GDP series data https://www.bea.gov/industry/xls/io-annual/GDPbyInd_VA_1947-2017.xlsx, access may 22, 2018. These values were tested against IBM total revenue numbers from IBM annual reports for each year, provided by IBM Archivist. The BEA values for "Publishing industries, except internet (includes software)", likely much less electronic in nature in those years, were excluded for years before 1990 based upon the IBM revenue numbers.

[7] Computing Costs Data uses hard disk storage costs analysis as a surrogate for total computing costs. Data was obtained from Matthew Komorowski's website: "A history of storage cost article updated September 8, 2009 (2014 update)". Accessed May 22, 2018: http://www.mkomo.com/cost-per-gigabyte. Komoroswki's sources are not public, but this series of data is used in many such analyses. His sample of disk costs from 1980 to 2014 shows that costs decline based upon this algorithm: =POWER(10,-0.2502*(year-1980)+6.304) with an r squared correlation of 0.9916. The graph includes the cost for the first entry in his table for each year (rather than an average). To extend this data to 1964, the graph includes data from IBM press releases from the IBM Archivist from 1964, 1966, 1969, not adjusted for inflation. Additional years were reviewed for 1968, 1972, and 1976, but the cost seemed unusually high, and likely included CPU and other system costs which could not be eliminated from the available sources. Reviewing Komoroski's 1980 and 1987 numbers, one notes that the IBM documents disk cost for 1980 at $173,626 per gigabyte of storage, and average for 1987 of $106,135, compared to Komoroski's cost of $193,000 and $90,000 respectively give some sense of validity to his underlying data, and a reasonableness for the trend. Note that years before 1980 are not predicted by Komoroski's algorithm which estimated a costs of $20 million in 1964 rather than $6.6 million. Although Moore's law was proposed in 1965, the impact upon disk costs may have been slower in starting.

[8] Twitchell, K. "A Proposed Approach to Minimum Costs for Financial System Data Maintenance." *Ledgerlearning.com*. April 2018, https://ledgerlearning.com/whitepapers/minimum-costs-for-financial-system-data-maintenance/

[9] James, W. (1890). *The principles of psychology*, Vol. 1. Henry Holt and Co. https://doi.org/10.1037/10538-000 as quoted in Boorstin, D. J. 1., & Luce, C. B. (1983). *The discoverers*. New York, Random House p. 718

[10] Twitchell, K. *Metric Engine: Reinventing the Data Supply Chains for Business,* monograph, x + 119 pp., Illus., index. Delta Summit Imprint, 2015. See also Twitchell, K. *Balancing Act: A Practical Approach to Business Event Based Insights,* xii + 348 pp. Illus., index. Delta Summit Imprint, 2011, 2015, p. 75-78, 321-324.

[11] Cripps, J. A *Contemporary Interpretation: Fra Luca Pacioli's 1494 "Particulars de Computis et Scripturis"* Pacioli Society, 1994, Seattle Washington http://jeremycripps.com/docs/Summa.pdf (Accessed Sept. 2022). See also Twitchell, K. "Historical Bookkeeping" *Conversations with Kip* YouTube Channel, April 2018, https://ledgerlearning.com/tag/historical-bookkeeping/?order=asc# (Accessed Sept. 2022)