Many is Beautiful:
Commoditization as a Source of Disruptive Innovation

By

Richard Ellert Willey

BA Economics / Government
Wesleyan University, 1988

MA Economics
Indiana University, 1993

SUBMITTED TO THE SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN THE MANAGEMENT OF TECHNOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2003

Signature of Author:____
Sloan School of Management
May 12th, 2003

Certified by:____
James M. Utterback
David J. McGrath jr. Professor of Management and Innovation
Thesis Supervisor

Accepted by:____
David Weber
Director, MIT Management of Technology Program

Many is Beautiful: Commoditization as a Source of Disruptive Innovation

Richard Ellert Willey

Submitted to the Sloan School of Management on May 12th 2003
in Partial Fulfilment of the Requirements for the
Degree of Master of Science in the Management of Technology

ABSTRACT

The expression “disruptive technology” is now firmly embedded in the modern business lexicon. The mental model summarized by this concise phrase has great explanatory power for ex-post analysis of many revolutionary changes in business. Unfortunately, this paradigm can rarely be applied prescriptively. The classic formulation of a “disruptive technology” sheds little light on potential sources of innovation. This thesis seeks to extend this analysis by suggesting that many important disruptive technologies arise from commodities. The sudden availability of a high performance factor input at a low price often enables innovation in adjacent market segments.

The thesis suggests main five reasons that commodities spur innovation:

- The emergence of a commodity collapses competition to the single dimension of price. Sudden changes in factor prices create new opportunities for supply driven innovation. Low prices enable innovators to substitute quantity for quality.
- The price / performance curve of a commodity creates an attractor that promotes demand aggregation.
- Commodities emerge after the establishment of a dominant design. Commodities have defined and stable interfaces. Well developed tool sets and experienced developer communities are available to work with commodities, decreasing the price of experimentation.
- Distributed architectures based on large number of simple, redundant components offer more predictable performance. Systems based on a small number of high performance components will have a higher standard deviation for uptime than high granularity systems based on large numbers of low power components.
- Distributed architectures are much more flexible than low granularity systems. Large integrated facilities often provide cost advantages when operating at the Minimum Efficient Scale of production. However, distributed architectures that can efficiently change production levels over time may be a superior solution based on the ability to adapt to changing market demand patterns.

The evolution of third generation bus architectures in personal computers provides a comprehensive example of commodity based disruption, incorporating all five forces.

Thesis Supervisor: James M. Utterback
Title: David J. McGrath jr. Professor of Management and Innovation

Table of Contents

Acknowledgements:

“Many is Beautiful” would never have been possible without the sustained support of my parents. Their generosity over the course of the last 12 months has truly been remarkable. I don’t say “thank you” nearly enough.

Many of the most important insights related to this thesis occurred as a result of a number of rambling discussions with my advisor, James Utterback. I was continually amazed by the breadth of his knowledge and the depth of his expertise.

Charles Lutwidge Dodgson has been a constant source of inspiration throughout my life.

Chapter 1

Studies focusing on disruptive technologies are almost always descriptive in nature. The destructive technology paradigm is well suited to explaining how different generations of disk drives manufacturers displaced one another or why mini-mills were able to undermine integrated steel plants. However, it is rare that this concept has been successfully applied proactively to accurately predict when an industrial sector is likely to be disrupted. Furthermore, this theory offers little explanatory power for the source of disruptive technologies. Rather, “The Innovator’s Dilemma” describes a set of market conditions that create a niche for a disruptive technology and then assumes that an appropriate technical solution will appear in the market.

“Many is Beautiful” proposes that the emergence of a commodity stimulates disruptive technology innovations in adjacent sectors of the economy. For example, the transition of corn starch into a commodity during the third quarter of the 19th century enabled the development of corn sweeteners and disrupted the sugar industry. This theme is inspired by two distinct, but previously unrelated academic traditions:

- Theories of “Supply Driven Innovation” focus on the difficulty of achieving an innovation. The monetary cost of the discovery process is typically used as a placeholder for difficulty.
- Computer science models use parallel construction to achieve robust design and fault tolerance.

Commodities are defined as undifferentiated products that compete primarily based on price. The marginal cost pricing scheme commonly associated with commodities supports and unifies these two analytic models.

- Changes in factor prices provide new opportunities for supply driven innovation.
- Low factor prices enable parallel construction: The efficiency of massively parallel systems is highly dependant on the cost of individual components.

1.1 Characteristics of Commodities

As a product transitions into a commodity, dramatic price based competition quickly drives prices down towards marginal cost. Various authors suggest different causes for the emergence of commodities. The product lifecycle model used by Utterback and Abernathy stresses that the transition towards price based competition is a natural part of the evolutionary process. Following the emergence of a dominant design, "firms turn their energies away from the innovation of product features and toward the innovations that lead them to cost or quality advantages on what has become a fairly standardized product."[1]

Clay Christensen extends this analysis by emphasizing the role that "performance overshoot" plays in creating commodities. Christensen hypothesizes that consumer's performance demands along any given dimension will eventually become satiated. As this happens, the locus of competition will shift to a new definition of performance. The competitive focus of an industry cascades through a series of different performance dimensions before settling into price based competition. **Diagram 1** is a Christensen illustration based on his analysis of the disk drive market in the mid to late 80s. In this market, competition initially focused on capacity, before transitioning to physical size, followed by reliability. Ultimately the market collapsed into price competition. Subsequent price wars lead to a dramatic collapse in profits, with gross margin's falling below 12%[2].

---

[1] James Utterback, Mastering the Dynamics of Innovation (Boston, MA: Harvard Business School Press, 1994), p. 34.

[2] Clayton M. Christensen, The Innovator's Dilemma (Boston, MA: Harvard Business School Press, 1997), p. 215.

Diagram 1 Changes in the Basis of Competition in the Disk Drive Industry[3]

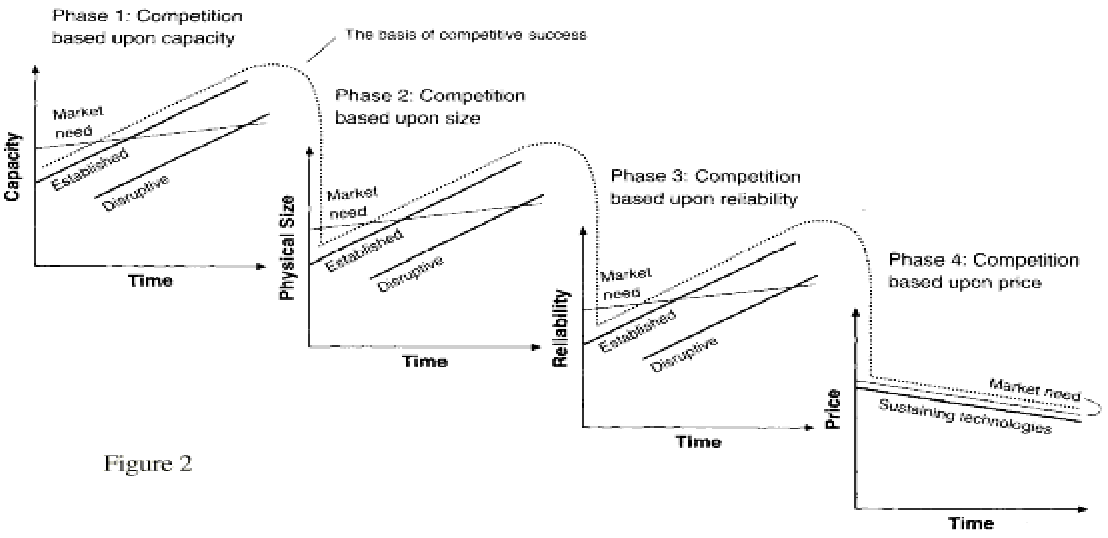


Figure 2

1.2 Supply Driven Innovation

In their seminal work "In Search of a Useful Theory of Innovation", Nelson and Winter distinguish between supply side and demand side innovation. Demand driven innovation occurs when

> changes in the composition of demand for goods and services across industries chain back to influence investment patterns, which, in turn, influence the relative return to inventors working on improvements in different kinds of machines[4]

Clay Christensen's technology overshoot model is a clear example of demand driven innovation. Christensen assumes that the existence of a market niche is sufficient to stimulate the innovation process.

In contrast, supply driven innovation can best be conceptualized as "searching" a specific technology base for potential innovations. Kaufman's academic work on N-K search (later extended by Fleming and Sorenson) provides useful intuition regarding this type of mental construct. These authors apply

---

[3] Ibid, p. 216

[4] R. Nelson and S. Winter, "In Search of a Useful Theory of Innovation," Research Policy 5 (1977): p. 50.

evolutionary biology models to the study of innovation, focusing on “a theory of innovation as a search process over technology landscapes.”[5]

> Many of the scholars who use evolutionary analogies propose that technological novelty arises from the recombination and synthesis of existing technologies… Thus, one can often describe inventions as a combination of prior and/or new technologies. For example, one might think of the automobile as a combination of the bicycle, the horse carriage, and the internal combustion engine.[6]

Each technology forms a dimension of an N-space. Evolutionary fitness is represented by the “height” of a surface passing through the space. Technological innovation is modeled identifying maxima across this surface.

Flemming and Sorenson use this model to explore the effects of modularity on search cost. I propose a dynamic extension to this model that would incorporate changes in search cost over time. Rational economic actors will invest an “optimal” amount of resources to search a given landscape, equating the marginal cost of research with the expected marginal revenue. Implicit in this is the assumption that, ceteris paribus, researchers will exhaust search opportunities on a static landscape and cease to invest additional resources. However, an event that distorts the search landscape will create new maxima and drive new opportunities for innovation. The same events that create new peaks may also depress existing maxima, eroding the market position of established products. Factor prices are [potentially] the most basic constraint in economics. “Many is Beautiful” proposes that the transition of individual goods into commodities triggers extreme changes in pricing, distorting the search landscape and enabling innovation.

---

[5] L. Fleming and O. Sorenson, “Technology as a complex adaptive system: evidence from patent data’” Research Policy 30 (2001): p.1020.

[6] Ibid

1.3 Many is Beautiful

"Many is Beautiful" defines five forces through which commodities enable innovation. Three of these forces are directly related to the cost of adopting commodities as factor inputs.

1. Low prices enable the substitution of quantity for quality.
2. The price / performance curve of a commodity creates an attractor that promotes demand aggregation.
3. Commodities emerge after the establishment of a dominant design. Commodities have defined and stable interfaces. Well developed tool sets and experienced developer communities are available to work with commodities, further decreasing the price of experimentation.

The last two forces focus on second order effects related to the granularity of a technology. Commodity components enable developers to substitute redundant parallel grids for large integrated systems.

4. Distributed architectures based on large number of simple, redundant components offer more predictable performance. Ceteris paribus, low granularity systems based on a small number of high performance components will have a higher standard deviation for uptime than high granularity systems based on large numbers of low power components.
5. Distributed architectures are often much more flexible than low granularity systems. Large integrated facilities often provide cost advantages when operating at the Minimum Efficient Scale of production. However, distributed architectures that can efficiently change production levels over time may be a superior solution based on the ability to adapt to changing market demand patterns.

Many is Beautiful concludes by studying "third generation" bus standards in personal computers. Intel's PCI Express[7] initiative provides a comprehensive example of commodity based disruption, incorporating elements from each of these five forces to create a disruptive innovation that will have revolutionary impact on enterprise computing.

---

[7] PCI Express is the follow on to Intel's Peripheral Component Interconnect (PCI) bus architecture for personal computers.

Chapter 2

Chapter 2 of “Many is Beautiful” offers three explanations how commodity cost structures create opportunities for supply driven innovation. The chapter starts by studying the potential to substitute quantity for quality. This is the simplest of the three forces and provides a useful introduction to the theme of commodity based disruption. Next, I identify the possibility that commodities act as “attractors”, aggregating demand within a market segment. This force benefits from a series of positive feedback loops and is potentially the most powerful of the five forces proposed in this thesis. The relationship between commodities and dominant design is the last force to be examined. Here the central argument focuses on the cost of experimenting with commodities.

2.1 Quantity versus Quality substitution

The potential to substitute quantity for quality is the simplest example of commodity based disruption. Quantity versus quality substitution is a broadly theme that can be applied in a wide variety of fields. The performance of a physical product is a function of the inputs that comprise it. For example, the strength of a rope can be modeled as a function of the number of strands of fiber as well as the composition of each strand. A strong rope can be assembled using a relatively small number of very strong strands or a relatively large number of weak strands.

The performance of an analytic technique can often be modeled as a function of the number of times that a step in the process can be repeated. Utterback compares mechanical balances with electronic scales and notes that electronic balances achieve superior accuracy by repeatedly measuring the same sample. As a more extreme example, so-called Genetic Algorithms combine random mutations across large numbers of generations to accomplish complex optimizations.

Economists have well defined models that they use to study quantity versus quality optimizations. Studying one such "standard" model provides useful insight regarding this type of disruption. Assume the following

1. A firm uses two inputs (X and Y) to produce and output.
2. Factor input Y is fixed with respect to price and quality.
   Let P_Y = Price of good Y
3. The quality of good X is allowed to vary.
   Let X = the quantity of good X consumed
   Let q_x = the quality of good X
   Let P_q = the price of a unit of X of quality q
   Let P_x = the quality adjusted price of X = P_q/q
4. The firm's production function is a homothetic function of X, q_x, and Y of the form Production = $F(X^* q_x, Y)$

Solve for the optimal X, q_x,Y subject to a fixed budget constraint

$$\text{Max } F(X^* q_x,Y) \text{ subject to } (P_q{}^*X) + (P_Y{}^*Y) = M$$

It can be demonstrated that in equilibrium

$$F_x(X_q,Y) / F_Y(X_q,Y) = P_x / P_Y = P_x / q^* P_Y$$

This is the economist's "standard" result that a rational producer will chose to purchase factor inputs X and Y to the level at which the ratio of the marginal productivity of is equal to the ratio of the factor costs.

In turn, we can then solve for the optimal quality level for X. Here, the equilibrium solution is that the rational producer will chose to invest in the quality of X that provides the cheapest quality adjusted price. More formally, the producer will choose the q_x that minimizes p_q / q_x.[8]

---

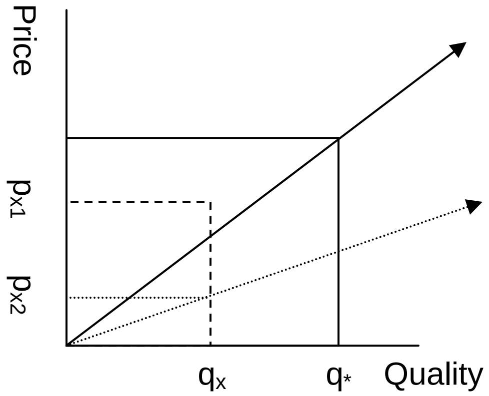
[8] http://www.src.uchicago.edu/users/gsb1/Econ301/Psets2002/pset2/pset2%232_ans.pdf

Retreating from calculus, back to the world of disruptive technology innovations, this basic model provides use with some useful insight into the conditions that must be fulfilled in order to enable disruptive innovation based on substituting quality versus quantity. The commodity based disruption model is based on changes in factor prices. In this case, we are interested in examples in which there is a sudden decrease in the price of a potential factor input.

1. Let p_* / q_* be the price:quality ratio of the incumbent technology.
2. Let p_{x1} / q_{x1} be the price:quality ratio before the potential factor input is reduced to a commodity
3. Let p_{x2} / q_{x1} be the price:quality ratio after the potential factor input is reduced to a commodity

$p_{x2} / q_{x1} < p_* / q_*$ is a necessary, but not sufficient condition for commodity based disruption. Graphically, this can be represented by **Diagram 2**.

Diagram 2 Quantity Quality Substitution



Initially, the factor input with price p_* and quality q_* is being consumed. The price:quality ratio for this factor input is represented by a vector from the origin, running through the point (p_*,q_*). This vector represents the market equilibrium. Consumers are indifferent to any price / quality combination that falls along this vector.

The point (p_{x1}, q_x) represents a good in the economy. Initially, this price / quality combination is not viable for the good to be used as a factor input in this market segment. Assume that this good transitions to a commodity, reducing its price from p_{x1} to p_{x2}. The new price:quality ratio for the commodity lies below the original vector, creating the potential for a commodity based disruption.

It is equally important to determine whether investing in the disruptive technology will yield a positive net present value. Modifying a platform to adopt a new factor input is not without cost. It is also necessary that the new factor input creates sufficient cost savings to compensate for the fixed cost of the necessary investment.

2.2 Commodities, "Attractors", and Demand Aggregation

An "attractor" is a concept commonly used in branches of applied mathematics such as physics and meteorology. An attractor is a defined as "a limit set that collects trajectories."[9] Magnetic fields are often used as an example of an attractor. A "magnetic field operates on [a] piece of iron, so that the latter strives to move towards the magnet."[10] This thesis proposes that commodities act in an analogous fashion within a market segment.

Assume the existence of a multi-dimensional preference space. Products and consumer demands are both represented by points in this space. Misalignments between consumer demand and products lead to a "Build versus Buy" decision. Consumers need to determine whether it is better to expend resources on internal research and development or to purchase existing components from outside the firm. Conceptually, we can consider four possible outcomes to the build versus buy decision:

---

[9] Tsonis, Anastasios A. Chaos: From Theory to Applications. (New York: Plenum Press. 1992)

[10] Einstein, Albert. Relativity: The Special and General Theory. (1920), Chapter 19.

1. There is no solution. [The project is cancelled]
2. The consumer uses internal resources to build a custom component. This activity creates a new product in the preference space.
3. The consumer decides to purchase a Commercial Off The Shelf [COTS] product as a "tactical" move. COTS is the best solution for this generation of product, however, this decision does not cause a fundamental re-alignment in the consumer's demand preference for this factor input. In this case, the "state" of the system has not been changed.
4. The consumer makes a strategic decision to migrate to a COTS input. The consumer adopts COTS in this generation of technology. In doing so, the consumer customizes his platform such that there is a permanent shift in his preferences.

In the case of the fourth outcome, the COTS factor input is acting as an attractor within the preference space; creating a permanent shift in consumer preferences for this factor input.

Validating this hypothesis requires identifying the logical equivalent of a magnetic field operating within the preference space[11]. Commodities are characterized by attractive prices and steady/sustained improvements in performance. Naively, we might assume that unit cost or performance is equivalent to magnetism. However, on inspection, it seems clear that these are both dimensions within the product space. Commodity's marginal cost pricing structure ensures that these products will be located close to the consumer demand points within the preference space, however, this is not the same as an attractor.

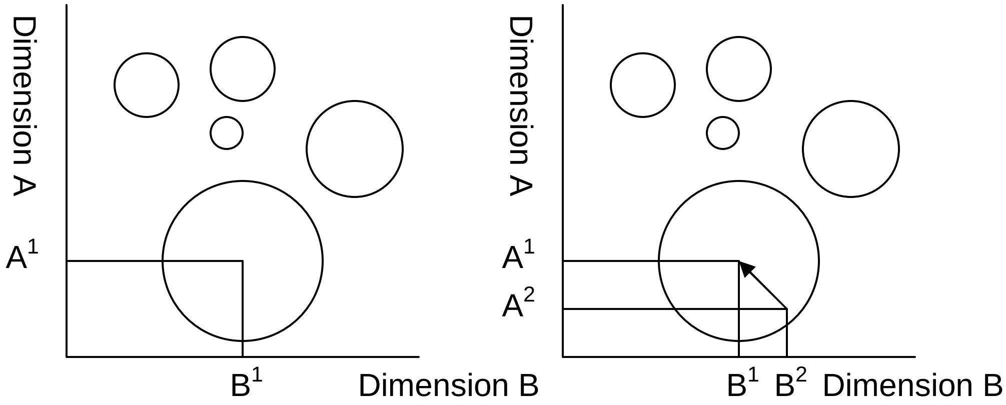
In the commodity based disruption model, the attractor operates through long run profit maximization. A permanent shift in consumer preferences requires that a "Buy" decision creates sufficient cost savings to justify the investment necessary to modify the product's architecture. Profit is a function of the relative locations of "consumer demand" and the commodity within the preference space, with the cost of modifying the

---

[11] As will be seen shortly, gravitation may be a superior example.

platform as the third required input for the optimization. This argument is illustrated in **Diagram 3**.

Diagram 3 Commodities as Attractors



The left hand diagram is a two dimensional representation of the consumer preference space. The diagram contains a series of spheres. Each sphere represents a product. The center of the sphere is the location of the product in the preference space. The radius of the sphere represents the area over which the product is able to act as an attractor[12]. The right hand diagram shows a new consumer demand point appearing in the economy. In the absence of existing products in this market, consumer demand would be located at point (A^2, B^2). However, the existence of a product at (A^1, B^1) causes the consumer to modify his platform to adopt (A^1, B^1) as a factor input. This creates a permanent shift in consumer preferences. “Many is Beautiful” defines this as “demand aggregation”.

The most interesting feature of this model is the logical parallel with basic models from astronomy. Astronomers have modeled how the gravitational contractions of diffuse gas clouds in space create solar systems with stars and planets. The self organizing nature of these models relies on the well known relationship that gravity is a

---

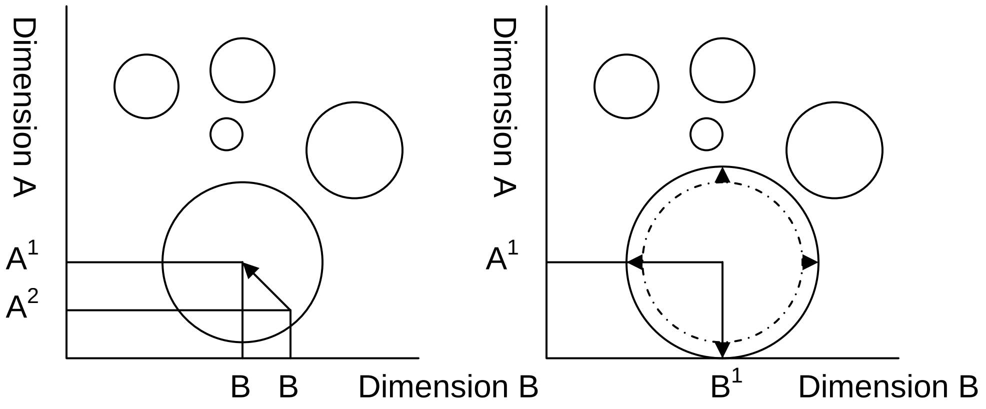
[12] As noted earlier, the radius of each sphere is a function of the cost of modifying the platform. As a result, this diagram is specific to a given technology.

function of mass. This creates a positive feedback loop in which the accretion of mass onto a body increases the gravitational attraction of that body. In a similar fashion, the price and performance curves for a commodity are a function of the cumulative demand.

- The fixed costs associated with research and development and production can be spread across a larger number of units, decreasing unit cost.
- Experience curve effects suggest that unit cost will decrease as product volume increases.

As the price of a commodity decreases and its performance increases, the distance over which it is able to act as an attractor increases. This dynamic is illustrated in **Diagram 4**.

Diagram 4 Demand Aggregation Increases the Strength of an Attractor

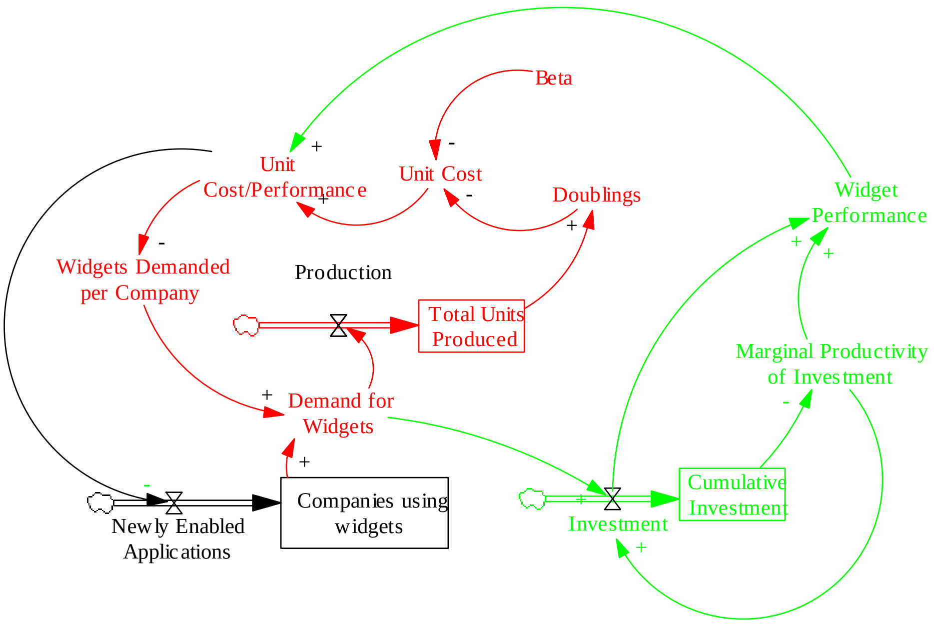


A Sect

2.3 A Simple Model of Commodity Based Disruption

Figure 5 provides a simple Systems Dynamics model designed to illustrate the positive feedback loops driving the Demand Aggregation model. The systems dynamics model postulates that commodity based innovation is driven by three primary reinforcing loops.

Figure 5



The first reinforcing loop which forms the core of the model is based on the concept that there is a relationship between the between the cumulative production and unit cost. This reinforcing loop - illustrated in red - is a simple representation of traditional "Experience Curve" arguments. The effect of this reinforcing loop is dampened of time by the fact that the number of Doublings is derived from a geometric series. ["Doublings" is defined as 1 + (ln(Total Units Produced)/ln(2)).][13]

The second reinforcing loop (illustrated in black) extends this model by hypothesizing that decreases in the unit cost enable a technology to diffuse into new market segments. This loop represents the argument that commodities serve as demand aggregators. A decrease in the cost of a commodity increases the number of companies that will chose to adopt the commodity as a factor input.

---

[13] Note another similarity to the gravitational models. In astronomy models, the inverse square law prevents systems from collapsing down into a single point source. The geometric series that underlies the experience curve operates in much the same manner.

The last section (illustrated in green) defines the relationship between the Demand for Widgets and investment in improving the performance of Widgets. This module includes both a reinforcing loop and an associated balancing loop. The reinforcing loop hypothesizes a very simply relationship in which Demand for Widgets results in Investment in improving Widget technology. Improved widget performance triggers an increase in the Demand for Widgets. The associated balancing loop reflects the fact that platforms become exhausted over time. Models of disruptive technologies typically assume Diminishing Returns to Scale from investing in a fixed platform. This is built into the model by assuming that there is an inverse relationship between cumulative investment and the Marginal Productivity of Investment.

The diffusion of Ethernet technology provides an outstanding example of this type of market dynamics. Ethernet Local Area Networks were originally defined by a broadcast bus wiring system and Carrier Sense Multiple Access / Collision Detect (CSMA/CD) technique for Media Access Control. However, as Ethernet matured, the wiring standard was broadened to include hub and spoke architectures like 10BaseT as well as point-to-point switched standards. Today, the Ethernet dominant design is defined by:

- An eight byte frame preamble for clock synchronization
- A 14 byte frame header incorporating a six byte Destination Address, a six byte Source Address, and a two byte protocol ID
- A four byte Cyclical Redundancy Check to verify data integrity
- A 1514 byte maximum frame size
- CDMA/CD

During the early 1990s, Ethernet emerged as the dominant data networking standard for Local Area Networks (LANs). Ethernet rapidly transitioned into a commodity, with dramatic effects on both pricing and performance.

This market trend has resulted in Ethernet technologies diffusing into a series of non-traditional markets ranging from PC bus architectures to the Internet core. As one specific example, Ethernet quickly jumped into wireless data networking with the development of the 802.11b wireless Ethernet standard. This technology was originally used to create wireless Local Area Networks (LANs). From this stronghold, 802.11b advanced as an alternative to Wideband Code Division Multiple Access (WCDMA) standard for bulk data transfer in 3G wireless networks. In the latest development, the steady decrease in the price of 802.11b has lead to the development of an 802.11b wireless cell phone. Cisco announced a new line of wireless cell phones in April 2003. These cell phones replace traditional voice networking access control standards such as Time Division Multiple Access (TDMA) and CDMA with 802.11b.

Section 2.4 Established Dominant Design

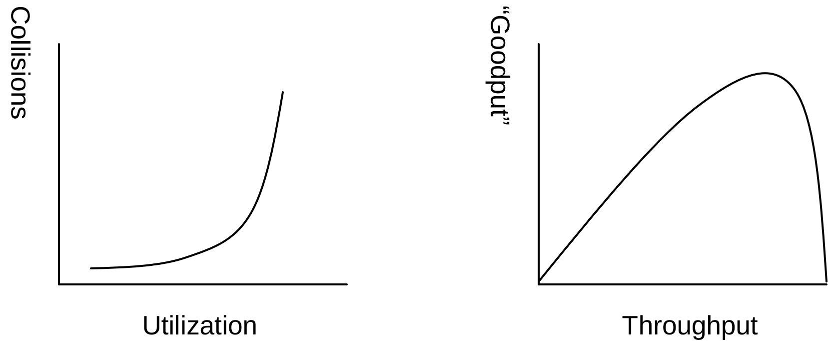
During the discussions of Quantity versus Quality substitution and Demand Aggregation, I noted that modifying a platform to switch to a new factor input requires investment. The greater this switching cost (equivalently, the larger the fixed cost required to enter a market), the less likely it is that a decrease in a factor price will act as a disruptive technology.

Commodities emerge after the establishment of a dominant design. Commodities are characterized by well defined and stable interfaces. Well developed tool sets and experienced developer communities are available to work with commodities. Each of these factors significantly decreases the fixed cost required to utilize a commodity as a factor of production.

Section 2.5 "Big Dumb Pipes" as an Example of Cost Driven Technology Disruption

The contrast between Managed Bandwidth and "Big Dumb Pipes" as guiding principles for network design provides a dramatic example of cost driven technology disruption. The performance of a network is characterized using a set of standard metrics including throughput, latency, "jitter" (the standard deviation of latency), packet loss patterns, and sequencing errors. Network performance is inversely related to the quantity of data being transmitting. The relationship is typically modeled using a "J-curve." Feedback loops degrade performance abruptly as load approaches a critical threshold.[14] Diagram 4 provides two simple graphics that describe congestion collapse on an Ethernet. Collisions cause the system to collapses as utilization approaches a critical value.

Diagram 6



Customers often need to associate different performance characteristics with different traffic flows on a network. Time sensitive data like an industrial control process or a real time audio-video stream requires a higher Quality of Service (QoS) than bulk data transfers. Managed Bandwidth implementations support this requirement by embedding high performance packet classifiers in routers and switches. These systems sort data into high priority and low priority streams in real time. High priority traffic is

---

[14] "RFC 2914: Congestion Control Principles", S. Floyd. 2000

forwarded immediately with near deterministic performance. Low priority traffic is buffered until the system has surplus bandwidth and then forwarded. During periods with a high load, low priority traffic may need to be discarded by the switch and then retransmitted by the end system.[15]

"Big dumb pipes" is a radically different design philosophy that eschews traffic classification. Proponents of this school implement devices that stream bits as quickly as possible from one interface to another. Packet "classification" is limited to inspecting the destination address in the packet to make a forwarding decision. Quality of Service is provided by deliberately overbuilding networking capacity to the extent that a traffic "burst" can not push total utilization into the steep portions of system's "J curve".

Big dumb pipes implementations can compete effectively with managed bandwidth systems because big dumb pipes implementations can be deployed at a fraction of the cost of a managed bandwidth solution offering similar throughput[16]. A "dumb" router only needs to inspect a single field within an IP header. Furthermore, the IPv6 destination address is fixed length 16 byte word that is always located at the same offset from the start of the header. This implementation was deliberately selected to simplify hardware based forwarding. In contrast, managed bandwidth solutions need much more complex forwarding engines.

- A "smart" forwarding engine needs to distinguish between many different types of traffic. This requires the ability to inspect multiple fields at different locations within packets.

---

[15] Asynchronous Transfer Mode (ATM) is an example of a datalink layer managed bandwidth implementation. There are also industry initiatives such as IP QOS designed to support managed bandwidth at the Internet's Network layer

[16] Networks built using "dumb" Ethernet technologies would often deploy 10 times the bandwidth as a Token Ring, ATM, or FDDI solution.

- In many cases, “stateful” packet inspection is required. The state of packet A depends on parameters set in packet B. Stateful packet inspection requires the ability to cache information from multiple packets.
- Smart forwarding engines may require upgrade capabilities so that they can recognize new types of traffic.

Each of these features dramatically increases the complexity and cost of the system.

The contrasting fortunes of “dumb” packet switched networks compared to circuit switched networks built using managed bandwidth principles provides a dramatic illustration how this dynamic has played out in the market. Telecommunications carriers traditionally deployed circuit switched networks that delivered low volumes of constant bit rate data with extremely high quality requirements. In contrast, packet switching data networks were designed to transmit large quantities of information with extremely “bursty” usage patterns and relatively low quality assurance.

Data networks transitioned rapidly into commodities. Associated with this, the past 30 years has seen a phenomenal growth in the cumulative number of bytes carried over data networks. This difference in traffic volume caused the data networking industry to mature much more rapidly than the regulated monopolies that provided circuit switched systems. The performance characteristics of the Internet have now improved such that data networks are subsuming voice traffic. Circuit switched networks no longer offer sufficient performance advantages to justify their dramatically higher costs. “Voice over IP” systems that carry voice traffic over data networks have already seized the lucrative business market. Voice over IP is starting to penetrate the home market. Commodity data networks are displacing premium voice services from their core market.

Chapter 3

3.1 Defining Granularity

Chapter 3 of "Many is Beautiful" focuses on what I describe as second order effects resulting from the emergence of a commodity. These second order effects are related to the "granularity" of a technology. Granularity describes the scale of the production process. A highly granular production process produces a small quantity of output using small amounts of inputs. A low granularity process requires that a relatively large quantity of output be produced using large quantities of inputs. Granularity can be illustrated using an extremely simple example. Economists often model production as a "black box" function that transforms a set of inputs into a set of outputs. In this example, inputs A, B, and C are transformed into output Z.

Table 6 Comparing High Granularity and Low Granularity production functions

Relative Granularity	**Inputs**	**Output**
High	A + B + 2C	Z
Low	5A + 5B + 10C	5Z

The low granularity process requires five units of A, five units, five units of B, and 10 units of C as an input. This process produces 5 units of Z as an output. The high granularity process operates on a smaller scale. This process requires one unit of A, one unit of B, and two units of C as the input and produces one unit of Z as an output[17]. In this example, five of the high granularity systems can substitute for one low granularity system. However, the low granularity system can not be subdivided to replicate a single high granularity system.

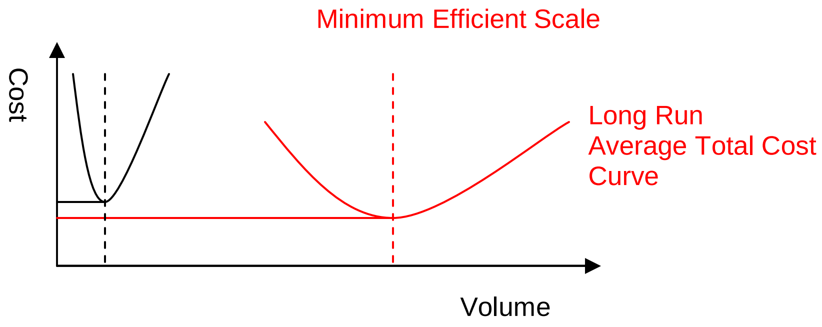
---

[17] Granularity is logically distinct from "Returns to Scale". Returns to Scale describe how production efficiency changes with production scale. For example, **Table 5** exhibits constant returns to scale since F(S*[A,B,C]) = S * F(A,B,C). Under increasing returns to scale F(S*[A,B,C]) > S * F(A,B,C).

3.2 The Effect of Commodity Based Disruption on Granularity

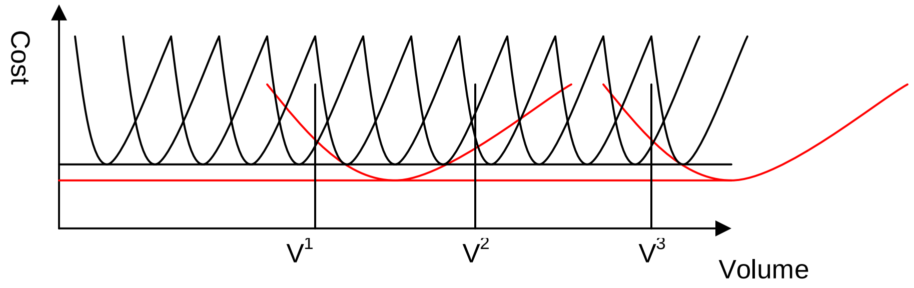
Commodity based disruption often has the effect of increasing the granularity of economic processes. The best explanation for this phenomenon is related to the earlier argument that commodities act as demand aggregators. Commodities are adopted as factor inputs in multiple market sectors. Different market segments have very different requirements regarding the appropriate granularity of an input to the production process. Low granularity products can be thought of as specialized solutions for specific niches. However, highly granular products can be flexibly applied in multiple market segments, offering increased sales opportunities. This dynamic is illustrated in **Diagram 7** using arguments derived from the Minimum Efficient Scale (MES) of production.

Figure 7 (a)



The production technology illustrated with the red curve is able to achieve lower Long Run Average Total Cost (LR-ATC) than the production technology shown with the black curve. In theory, this corresponds to a technology where a large integrated plant is able to utilize efficient mass processing technologies that can not be scaled down efficiently. However, **Figure 7b** clearly demonstrates that there are ranges of production volume at which large numbers of the smaller systems can achieve lower average total costs.

Figure 7b



The low granularity production technology minimizes cost when the production volume falls between [V^1,V^2], however, the high granularity solution achieves cost savings when demand falls outside this range.

This argument suggests that the “optimum” granularity of a product is a function of the distance [V^1,V^2] and the probability density function describing consumer demand for volume. It is important to note that distance [V^1,V^2] is a function of the specific industry cost curves. For example, shifting the black curve down by a constant amount will decrease [V^1,V^2]. Ceteris parebis, flattening the Long Run ATC for the high granularity technology will increase [V^1,V^2].

Increasing the granularity of an economic process spurs innovation in two ways: First, an increase in granularity creates new opportunities to achieve robust design through the use of redundant parallel systems. Second, an increase in granularity provides a more flexible system that is better able to adapt to variations in demand.

Section 3.3 Redundancy and Robust Design

Engineers have developed formal academic models to study robust design and fault tolerant systems. These models were developed to study when it is possible to construct high reliability systems using fallible components. One potential mechanism to build high reliability systems is improving the quality of individual components.

Component cost is typically treated as a function of reliability with the second derivative of cost with respect to reliability assumed to be > 0. Building individual components capable of “5-9s” reliability becomes prohibitively expense. Leveraging redundant systems and parallelism often proves to be a more cost effective solution. Formal study dates to Von Neumann’s “Probabilistic Logic and the Synthesis of Reliable Organisms from Unreliable Components”. Von Neumann demonstrated that a Triple Modular Redundant (TMR) system could be used to provide acceptable reliability at low cost[18].

Von Neumann identifies a set of critical issues that must be considered when designing parallel redundant systems. Von Neumann’s original results assumed that component failures can be treated as independent events. Positive correlation of component failure significantly degrades the cost efficiencies from parallel construction. Assume that a grid computer is assembled from a large number of processors. It is critical that the system be designed such that the failure of one processor does not trigger a cascade. In a similar fashion, the failure mode for multiple processors should not be correlated to individual events.

Von Neumann also used a simplifying assumption that the reliability of an individual components R_M can be modeled as a

> decaying exponential of the operating time.
>
> $$R_M(t) = \exp(-ft) = \exp(-t/MTF)$$
>
> In this formula, f is a constant called failure rate; and MTF is its reciprocal, called mean-time-to-failure. The reliability of the triple redundant system is now given by
>
> $$R(t) = 3 \exp(-2t/MTF) – 2(\exp(-3t/MTF)^{19}$$

---

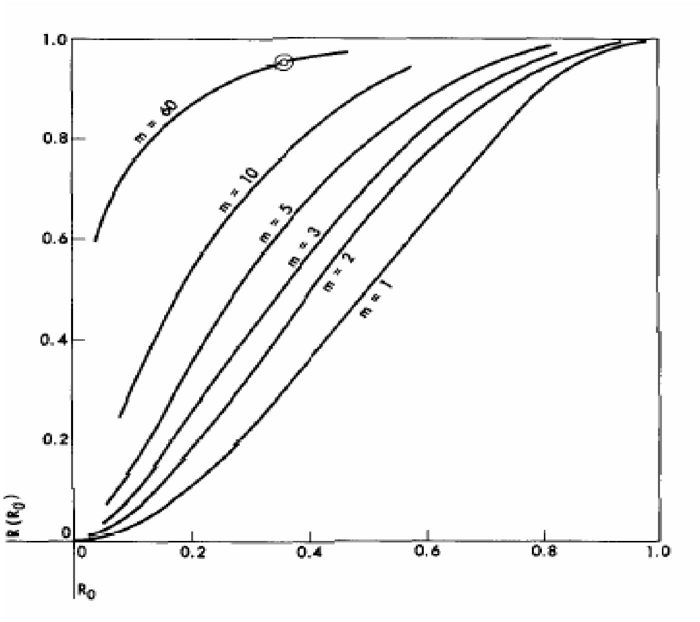
[18] John Von Neumann, “Probabilistic Logic and the Synthesis of Reliable Organisms from Unreliable Components” in Automata Studies, ed. C.E. Shannon and J. McCarthy (Princeton, Princeton University Press, 1956) p. 84.

[19] R.E.Lyons and W. Vanderkulk “The Use of Triple-Modular Redundancy to Improve Computer Reliability”, IBM Journal, April 1962. p. 2

Von Neumann makes an explicit assumption that the failure rate of individual components is constant in any given time period. This assumption is perfectly reasonable for hardware components like a computer chip that do not suffer “wear and tear”. However, the reliability of some system components does degrade over time. Moving parts erode. Poorly written applications cause memory leaks. Von Neumann assumes that optimal maintenance scheduling can be used to approximate constant failure rates over time by servicing components significantly in advance of the mean-time-to-failure.

Von Neumann’s critical result is summarized in **Diagram 8**. The X axis measures R_0; the reliability of an individual component. The Y axis measures $R(R_0)$; the reliability of a system built from a series of TMR components. $R(R_0)$ is defined as $R = (3\,R_0^{2/m} - 2R_0^{3/m})^m$, where m is equal to the number of redundant modules. **Diagram 8** defines a series of reliability curves that show system reliability as a function of the number of redundant components.

Diagram 8 TMR reliability R versus Non-Redundant Reliability R_0[20]



---

[20] Ibid, p. 3

Von Neumann's models were designed to describe systems with binary operating modes. A TRM system either works or fails; however, there is no concept of performance degradation. More recent work has focused on modeling graceful failure. Here, the processing power of the system is treated as a function of the number of redundant systems that are current operational. System optimization focuses on designing fault tolerant structures capable of providing bounded processing power. The most powerful approaches achieve reliability through systemic over-provisioning using commodity components.

Assume for the moment, that two different systems can be designed. The first features 5 large processors, each of which provides 20% of the aggregate processing power. The second system features 1000 small processors, each of which provides .1% of the aggregate processing power. During any period, there is a 1% chance that a given system component will fail. At any given point in time, we expect that the system with 1000 processors will have many more processors off line than its 10 processor counterpart. However, from the perspective of a business owner, minimizing the standard deviation of the available processing power available is likely to be more important than minimizing the number of hardware failures. This intuition can be formally represented as follows: Let

N = the number of processors in a system
X_{iT} = the percentage chance that processor i is working at time T
R_{ij} = The covariance between X_{iT} and X_{jT}
    $R_{ij} = 0$
$R_{T,T+1}$ = The covariance between X_{iT} and X_{iT+1}
    $R_{T,T+1} = 0$
Y_T = the percentage of processors that are working at time T
P = the total processing power of the system
    P is a linear transformation of Y]

The standard deviation of Y_T can be represented as.

$$\sigma = [\ X_{iT}(1-X_{iT}) / N\ ]^{.5}$$

The standard deviation of processing power is inversely related to the square root of the number of processors. While this analysis was presented as a study of failure modes, the results can be extended to support preventative maintenance. Maintenance service often requires taking sub-systems off line to upgrade or replace parts.

Steven Gribble has written multiple articles about achieving robust design through parallelism and redundancy. Like Von Neumann, Gribble stresses the importance of validating the basic assumptions of these models when designing operational systems. Gribble specifically notes that many real world implementations encounter unexpected problems. In particular, memory leaks often result in synchronized device failures. Gribble is more troubled by the observation that independent periodic signals in computer environments often self-synchronize. Gribble's intuition is summarized in **Diagram 9**.

Diagram 9[21]

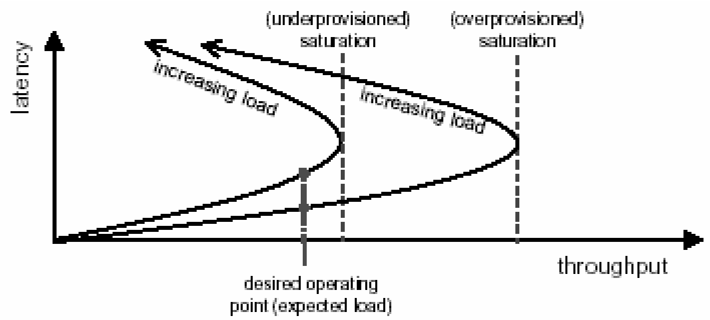


Figure 3. **An overprovisioned system:** by overprovisioning relative to the expected load, the system has slack: it can withstand unexpected bursts of load without falling into the "hole" associated with operating beyond saturation.

---

[21] Steven D. Gribble "Robustness in Complex Systems" http://www.cs.washington.edu/homes/gribble/papers/robust.pdf

3.4 Adapting to Demand Variability

“Many is Beautiful” postulates that commodities evolve towards high granularity system in order to provide flexible solutions capable of satisfying requirements from multiple distinct market segments. This same flexibility ensures that commodities can adapt to stochastic demand patterns within a single market segment. Consumer demand changes over time. Associated with this, optimal product quantities will also change. Here once again, high granularity systems offers superior production efficiencies.

As an example, Google has built one of the Internet’s most powerful data processing systems. Google’s data processing system consists of 54,000 servers supporting over 100,000 processors and approximately 216,000 hard drives. Google’s CEO, Eric Schmidt, has deliberately followed an expansion model based on organically adding capacity using commodity components. Commodity components provide Google with the ability to flexibly add incremental processing and storage capacity. Google has the option to flexibly <u>decrease</u> capacity by deferring maintenance. The performance of the system will degrade gracefully as components fail and are not replaced.[22]

---

[22] http://www.nytimes.com/2003/04/13/technology/13GOOG.html?pagewanted=all&position=top

Chapter 4

"Many is Beautiful" concludes by examining the evolution of third generation bus architectures in personal computers. PCI Express is the PCI Special Interest Group's official replacement for the PCI bus architecture. The PCI Express bus specification provides a comprehensive example of the dynamics described by "Many is Beautiful". Each of the five forces described in this thesis are evident in the architecture of the PCI Express bus. In particular:

- The PCI Express leverages commodity hardware technology derived from data networking switching systems.
- The core of PCI Express bus is composed of a high granularity mesh of redundant switching elements.
- TCP/IP[23] will provide a standardized interface for inter-device communication across the new bus.

"Many is Beautiful" predicts that this new architecture will act as a power disruptive force throughout the computer industry.

4.1 Requirements for a third generation bus architecture

The computer industry uses the word "bus" to refer to a circuit that connects the Central Processing Unit (CPU) with peripheral devices such as a hard drive, memory, a video card or a Network Interface Card (NIC). Recent advanced in bus designs allow multiple CPUs to be connected into a symmetric multiprocessing system. The PC bus provides a modular product architecture. Using a bus allows developers to replace individual components without the need to change other elements of the system.

"First Generation" bus architectures like the S-100[24] and ISA are distinguished by the fact that every device connected by the bus shares a common clock signal transmitted

---

[23] TCP/IP stands for Transmission Control Protocol / Internet Protocol. TCP/IP is the standard communications protocol used for communications between hosts on the Internet.

by the CPU. These architectures also require active participation by the CPU in any and all transactions. These two features eventually emerged as critical performance constraints forcing the development of second generation architectures. The common clocking requirement means that system performance was limited by clock speed of the slowest device on the system. The requirement that the CPU be involved in all transactions created a critical bottleneck on system performance[25].

"Second Generation" bus architectures such as Peripheral Component Interconnect (PCI) and NuBus were specifically designed to solve these problems. The processor uses the so-called Northbridge chip (also referred to as the host/bus controller) as a controller for memory and the PCI bus. The introduction of the Northbridge chip decoupled the speed of the processor from the speed of peripheral devices. Peripherals have the option to implement independent controllers and function as autonomous subsystems.

All other devices that need to communicate with the processor interface through the Southbridge chip. "Southbridge is a PCI-ISA bridge, which connects all other I/O subsystems *through* the PCI bus, and then through the Host-PCI Bridge on the Northbridge chip to the CPU.[26]" Southbridge isolates the processor from the majority of devices running on the system and successfully removed the CPU as a bottleneck.

However, like the first generation architectures that it supplanted, the PCI platform has architectural limitations that are forcing a platform upgrade. In second generation architectures the critical performance bottleneck is the speed of the PCI bus.

> Assuming we have only one 32-bit 33 MHz PCI bus on the motherboard … our maximum available data path between the Northbridge and the Southbridge chips is about 132 Mbits/s. What devices are connected to the

---

[24] Standardized as IEEE 696
[25] http://www.wikipedia.org/wiki/Computer_bus
[26] http://www.chipcenter.com/eexpert/dgilbert/dgilbert063.html

> Southbridge? *Everything*, aside from memory, PCI, and AGP. This means that traffic from all the IDE channels, the USB, the system BIOS, the ISA bus, the interrupt controller, and so on and so forth, must go through the Southbridge[27].

Furthermore, processor performance is increasing much more quickly than the performance speed of the internal bus.

> PC buses have doubled in performance roughly every three years, from the days of the original 8-bit PC/XT and 16-bit ISA buses, to 32-bit EISA and MCA, VL Bus, PCI, PCI-64/66MHz, and now PCI-X 1.0 and 2.0. But processors have roughly doubled in performance in half that time per Moore's Law. This schism creates bottlenecks -- the I/O channels aren't fast enough to keep the CPU/memory subsystem steadily fed with data[28].

4.2 The evolution of the PCI Express architecture

PCI Express was designed to solve performance bottlenecks in second generation bus architectures by leveraging commodity technologies developed within enterprise data networking. The developer’s primary inspiration was the potential to harness ongoing enhancement in the forwarding speed of enterprise switches. Ethernet developers have increased performance by a factor of 10 every four years. Wide Area Network (WAN) standards traditionally increase performance by a factor of 4 every 2.5 years. Both performance curves are substantially steeper than that of microprocessors.

Table 10 Performance Increases in Ethernet and WAN Technologies

LAN Technology	Year of Introduction	Performance (Mbps)	WAN Technology	Year of Introduction	Performance (Mbps)
Fast Ethernet	1995	100	**OC12**	1996	622.08
Gigabit Ethernet	1999	1000	**OC48**	1999	2,488
10 Gig Ethernet	2003	10,000	**OC192**	2001	10,000
			OC768	2003	40,000

Diagram 11 graphs normalized performance curves for the four different technologies over five years. This table demonstrates the evolving performance mismatch between

---

[27] http://www.chipcenter.com/eexpert/dgilbert/dgilbert063.html
[28] http://www.extremetech.com/article2/0,3973,522346,00.asp

processors and the PCI bus, as well as the potential to solve this with technologies derived from data networking.

Diagram 11 Normalized Performance over Time

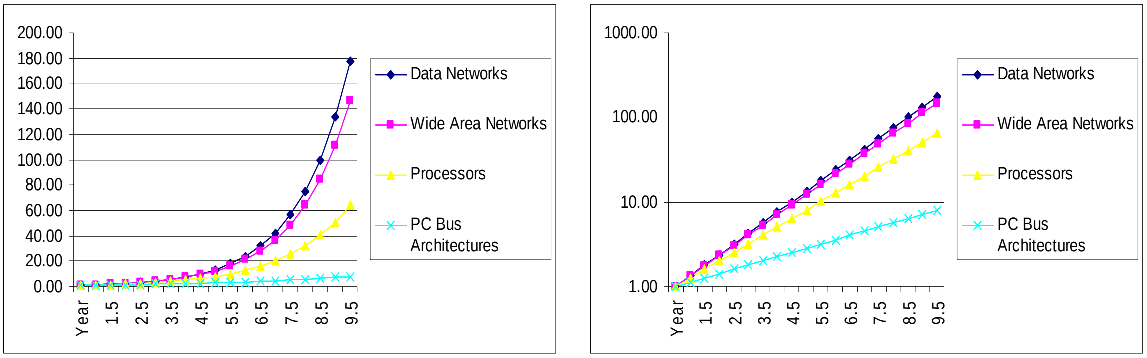
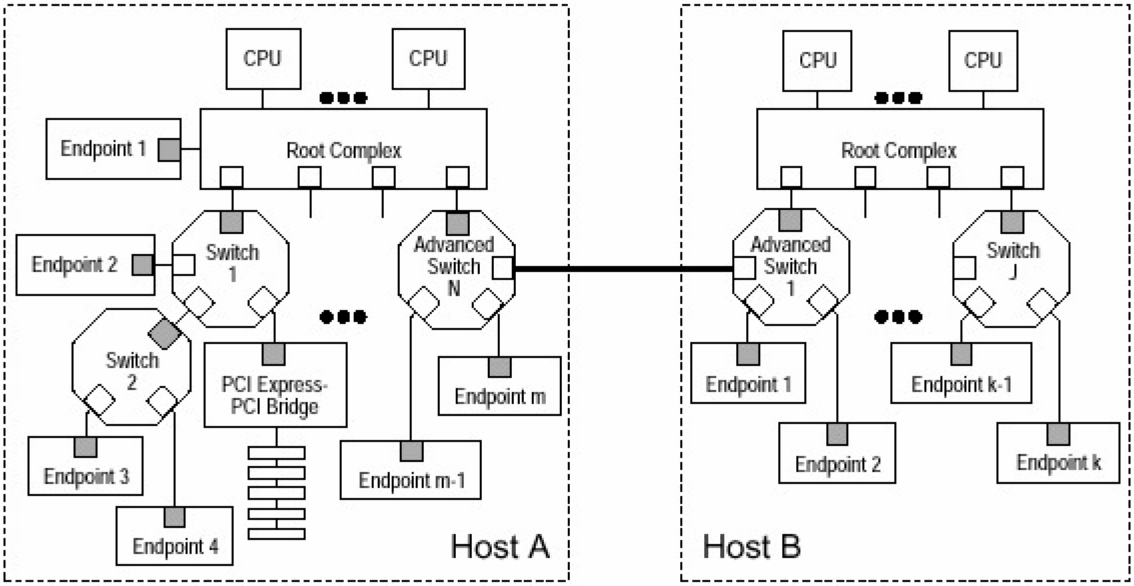


Figure 12 provides a graphical representation of the implementation of the PCI Express architecture. The data networking antecedents are readily apparent, down to the naming conventions used for individual components.

Figure 12 PCI Express Bus Architecture[29]



The Root complex is analogous to Northbridge and provides extremely high throughput / low latency connectivity between the processor and particularly "demanding" devices.

---

[29] http://www.extremetech.com/article2/0,3973,522348,00.asp

This example shows that multiple CPUs in a symmetric multiprocessing system will attach directly to the Root Complex.[30] The processor uses the Root complex to communicate directly with memory. It is also anticipated that the Advanced Graphics port will connect directly to the Root Complex.

All other communications travel across a "Switch". Switches provide interconnections to peripheral devices that do not require the same throughput / latency guarantees available across the Root Complex. Switches are high granularity, modular components that can be run in serial or parallel. The role of a switch is to direct data to the correct peripheral device based on an addressing field in the frame header. This graphic also shows a specialized component called an "Advanced Switch". Advanced Switches acts as routers, providing network layer functionality to link together peripheral devices located in different chassis. Advanced Switches can be directly connected to form a fully redundant and meshed switching fabric; the so-called "PCI Express Advanced Switch Fabric".

Intel claims that this technology scales almost linearly with the number of switches and that the fabric can be used to construct everything from relatively low end systems like a PC bus all the way up to the primary switching fabric for high-end Enterprise and Telecom routers. As demonstrated in **Diagram 13**, Intel white papers are already promoting this architecture as a replacement for the high end switching fabrics released by companies like Brocade and Juniper[31].

---

[30] http://www.extremetech.com/article2/0,3973,522348,00.asp
[31] http://www.intel.com/design/network/papers/25173701.pdf

Diagram 13 Intel Advanced Switching Fabric[32]

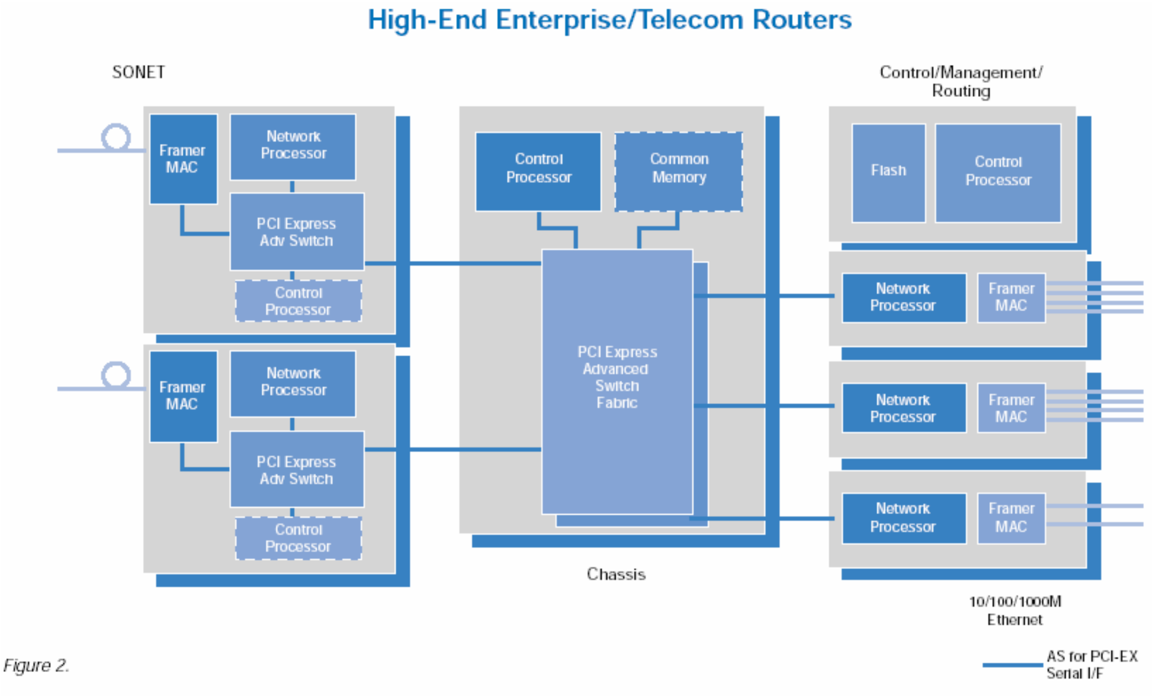


Figure 2.

Intel's original specification for the PCI bus implemented a proprietary set of data networking services. Intel specified a three layer protocol model that incorporates a physical layer, a protocol layer, and a transaction layer. Specific services include:

- Addressing/Routing services identify the location of specific peripherals on the switching backplane.
- Cyclical Redundancy Checks are used for error detection and provide some capabilities for error correction
- Sequence numbers are used to ensure the reliable transmission of data across the backplane.
- A credit based flow control system is used to govern the rate of data transmission[33].

---

[32] http://www.intel.com/design/network/papers/25173701.pdf
[33] http://www.extremetech.com/article2/0,3973,522348,00.asp

Intel's decision to implement these services using a proprietary standard was extremely curious since all of these services are already available within the TCP/IP protocol suite.

- Internet Protocol v6 provides source and destination addressing
- TCP provides a checksum for both header information and data
- TCP provide sequence numbers and reliable data transmission
- TCP provides flow control using a sliding window system

TCP/IP is the dominant data networking standard and will definitely be used to transfer application layer information over the PCI Express bus. Implementing redundant functionality at multiple layers within a protocol stack violates commonly accepted "end to end" principles of network design[34]. This decision also runs contrary to the "Many is Beautiful" assertion that the existence of standardized interfaces, tool sets, and active developer communities promotes commodity based disruption. "Many is Beautiful" suggests that Intel should have used the commodity TCP/IP standard to provide data networking services.

Intel's decision to adopt a proprietary standard suggests that Intel wants to use the interface to lock-in users. However, it is clear that the market is already working to modify the PCI Express standard, replacing Intel's proprietary data services with the commodity TCP/IP protocol suite. Many third generation bus standards that originally were in contention with PCI Express were explicitly designed to utilize the TCP/IP protocol suite for data services. For example, the Infiniband standard utilizes IPv6 for Network layer connectivity and TCP for transport layer functions[35]. Hypertransport specifies an extremely high speed / low latency bus, similar to the PCI Express root complex. Hypertransport permits users the flexibility to place any network or transport

---

[34] J.H. Saltzer, D.P. Reed, and D.D. Clark. "End-to-End Arguments in System Design." Web.mit.edu/saltzer/www/publications/endtoend/endtoend.pdf. p. 3.

[35] http://www.extremetech.com/article2/0,3973,390917,00.asp

layer services over this bus, anticipating the use of TCP/IP for upper layer services. Engineers are already working to port these functions from competing third generation standards to PCI Express.

Engineers working with low level operating system functions such Direct Memory Access (DMA) have recognized a requirement to layer their services on top of the dominant industry standards. For example, the Remote Direct Memory Access consortium has decided to layer RDMA over TCP/IP rather than using Intel's proprietary services[36]. The RDMA consortium specifically states that they adopted this strategy based on the inherent advantages from layering RDMA over an open and interoperable commodity standard.

Sustained market pressure is likely to force modifications to the PCI Express standard to permit the use of TCP/IP data services directly over the hardware switching grid. Standards traditionally ratify technologies that have already gained dominant position in the market. If the PCI SIG is unwilling to "officially" adopt the necessary changes, designers will simply innovate around the standards body and wait for the PCI SIG to "catch up".

4.3 PCI Express and Commodity Based Disruption

This short introduction to the PCI Express standard demonstrates that this architecture embodies Many is Beautiful's thesis of commodity based disruption. PCI Express leverages multiple dimensions of commodity based disruption.

- PCI Express has adopted commodity technologies derived from data networking.
- PCI Express has implemented this technology using a high granularity parallel backplane.
- PCI Express is migrating to the commodity TCP/IP interface.

---

[36] http://www.rdmaconsortium.org/home/The_Case_for_RDMA020531.pdf

It is almost impossible to overstate the potential impact of the PCI Express architecture on enterprise data networking. PCI Express will revolutionize a broad range of industrial sectors ranging from Network Interface Cards to Operating Systems.

Vendors selling forwarding systems such as Cisco, Brocade, Nortel, and Juniper currently provide integrated hardware and software solutions. The PCI Express architecture will substantially lower barriers to entry for hardware forwarding systems. New market entrants will quickly emerge and build commodity switching fabrics based on meshes of Advanced Switching modules. The companies that currently dominate this market sector will be forced to differentiate themselves on the quality of their software, as well as advanced services such as management and security.

PCI Express will have an equally dramatic effect on network interconnection models. With existing bus architectures, the Network Interface Card (NIC) serves as a bridge between the personal computer and the network. PCI Express deploys the same switching fabric for both the PC bus and the network switch. With the new architecture, linking a PC into a network switch transparently extends the PC bus into the Local Area Network (LAN). Local applications will be able to directly access data on a remote hard drive without the need to route requests through the CPU on the remote host. NICs will only be required to bridge dissimilar networks technologies. For example, a NIC would still be necessary to connect the advanced switching fabric to an 802.11 wireless Ethernet.

PCI Express will promote the development of so-called “thin clients”. Thin clients are modern versions of dumb terminals like DEC’s VT100. PCI Express decomposes a personal computer into a distributed system. Ultimately, this same decomposition eliminates any requirement that peripherals be located within a single physical chassis.

Next generation enterprise computing solutions will be designed to leverage the power and reliability of the new Processor Area Networks and Storage Area Networks. End user terminals will be designed using simple processors primarily intended for graphics display. RAM caches will substitute for "permanent" storage ensuring rapid response time for localized data. The "efficiency" of thin client architectures always collapses into a tradeoff between three variables:

1. The cost of centralized processing relative to local processing
2. The cost of centralized storage relative to local storage
3. The latency of the communications path

Express simultaneously decreases the cost of deploying centralized processing/storage solutions while improving the latency characteristics of the communications path; promoting a shift towards thin clients[37].

PCI Express' most important impact is likely to occur in the market for Operating Systems. Microsoft Windows evolved as a single user operating system designed to run on a single processor. Historically, the operating system had no concept of "users" or layered execution privileges, though the Windows NT family of Operating Systems has started to add rudimentary support for these features. It is doubtful whether the Windows can scale to this new distributed architecture. At the very least, Windows will need to be radically redesigned to focus on managing access privileges.

More radical theories suggest the distributed nature of enterprise computing will promote so-called "microkernel" architectures such as MACH and HURD[38]. Both UNIX and Windows utilize monolithic kernels based on a top down architecture in which all

---

[37] It is often proposed that the cost of administering distributed systems compared to centralized systems should also be built into this type of model. However, this cost spread is typically modeled as a constant and should not have a significant impact the dynamics of the system.

[38] MACH's history is documented at http://www-2.cs.cmu.edu/afs/cs/project/mach/public/www/mach.html
Ongoing work on HURD is documented at http://www.gnu.ai.mit.edu/software/hurd/docs.html

kernel functions execute within a single context. In contrast, microkernel architectures segment kernel functionality into a series of independent modules connected by well defined interfaces. The operating “system” is assembled from the bottom up using a set of modules. This type of bottom up, self organizing approach is much more appropriate for developing operating systems composed of large sets of autonomous components.

Chapter 5: Conclusions

Many is Beautiful's theory of commodity based disruption offers a powerful explanation for the emergence of disruptive technologies. Clearly, commodity based disruption is not meant to be an all encompassing theory; However, I believe that this concept provides important insights regarding the innovation process. Mankind has known how to produce steel for thousands of years; however, the Age of Steel did not truly being until steel became cheap and plentiful. In a similar fashion, the true "computing revolution" is only now occurring, as processing, bandwidth, and storage are ultimately being reduced into commodities.

Directions for Future Research

This thesis can be extended in a variety of ways. The most crucial area for future research is a more comprehensive econometric analysis of the general themes and models proposed by this thesis. In particular, three topics obviously lend themselves to future research: "Many is Beautiful" proposes a very simple systems dynamics model that describes many important elements of commodity based disruption. "Fitting" this model to a specific set of parameters is a critical step in validating this hypothesis. As noted earlier, I believe that the diffusion of Ethernet may be the best example to consider.

After validating the systems dynamics model operates as conjectured, the next avenue for exploration will be to develop formal mathematical models of demand aggregation. As mentioned earlier, the astronomical models used to study the coalescence of gas clouds is an obvious starting point. Here, once again, the basic goal is to develop a model and then to validate this using market data or experimentation. Ideally, it might be possible to derive the business equivalent to Jean's Criteria[39].

It is possible to apply formal models to describe the trade offs inherent in different schools of network design. Ideally, it might be possible to characterize a set of boundary conditions that indicate when Big Dumb Pipes work better than Managed Bandwidth approaches. Too often, arguments over network design principles are reduced to "religious" wars. It would be extremely useful to adopt more rigorous modeling techniques.

At the opposite extreme, commodity based disruption is not limited to the high tech sector. It would be extremely interesting to extend this same theme to industrial goods. A detailed analysis of the steel industry or corn refining might provide valuable new insights on this general theme. Researchers might all consider a broad-based approach that studied multiple market segments searching for additional ways in which commodities spur innovation. "Many is Beautiful" identified five forces through which commodities spur innovation. It is entirely possible that additional forces can still be identified.

---

[39] Jean's Criteria characterizes a necessary and sufficient condition for the collapse of a diffuse gas cloud into a star. Jean's Criteria establishes a relationship between the kinetic (thermal) energy of the system and gravity.

Bibliography

Afuah, Allan and Utterback, James. “The Emergence of a New Supercomputer Architecture” Technology Forecasting and Social Change 40, 1991.

Abernathy, W.J. and Clark, K.B., “Mapping the Winds of Creative Destruction.” Research Policy 14, 1985.

Christensen, C.M. The Innovator’s Dilemma. Boston, MA: Harvard Business School Press, 2002.

Christensen, C.M., Suárez, F.F. and Utterback, J.M. “Strategies for Survival in Fast-Changing Industries”, Management Science, 44 (1988).

Christensen, Verlinden, and Westerman, “Disruption, Disintegration, and the Dissipation of Differentiability.” Industrial and Corporate Change 11 (2002).

Gawer, Annabelle and Cusumano, Michael. Platform Leadership: How Intel, Cisco, and Microsoft Drive Industry Innovation. Boston, MA: Harvard Business School Press, 2002.

Gilbert, David. The NVIDIA nForce Chip Set. Published in the ChipCenter Questlink http://www.chipcenter.com/eexpert/dgilbert/dgilbert063.html

Gribble, Steven. “Robustness in Complex Systems” http://www.cs.washington.edu/homes/gribble/papers/robust.pdf

Fleming, Lee and Sorenson, O. “Technology as a complex adaptive system: evidence from patent data.” Research Policy 30 (2001)

Floyd, S. “RFC 2914: Congestion Control Principles” 2000

Henderson, R. and Clark, K. “Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms.” Administrative Science Quarterly 35, (1990)

Intel Corporation. “Advanced Switching in the PCI Express Architecture: The Unifed Interface for Both Control and Data Planes in a Converged Computing & Telecommunications Environment” http://www.intel.com/design/network/papers/25173701.pdf

Langlois, R. and Robertson, P. “Networks and innovation in a modular system: lessons from the Microcomputer and stereo components industries.” Research Policy 21 (1992).

Lyons, R.E. and Vanderkulk, W. “The Use of Triple-Modular Redundancy to Improve Computer Reliability”, IBM Journal, April 1962.

Meeus, M. and Oerlemans, L. “Firm behavior and innovation performance – An empirical exploration of the selection-adaption debate.” Research Policy 29 (2001).

Nelson, R. “Research on Productivity Growth and Productivity Differences: Dead Ends and New Departures.” Journal of Economic Literature 19 (1981)

Nelson, R. and Winter, S. “In Search of a Useful Theory of Innovation.” Research Policy 5 (1977).

Pavitt, Keith. “Sectoral patterns of technical change: Towards a taxonomy and a Theory.” Research Policy 13 (1984).

Pinkerton, Jim. The Case for RDMA. Published by the RDMA Consortium, 2002 http://www.rdmaconsortium.org/home/The_Case_for_RDMA020531.pdf

Saltzer, J.H., Reed, D.P and Clark, D.D. “End-to-End Arguments in System Design.” web.mit.edu/saltzer/www/publications/endtoend/endtoend.pdf

Scherer, F. “Inter-Industry Technology Flows in the United States.” Research Policy 11 (1982).

Stam, Nick. Buses and Interconnects: Where From Here? Published by Extreme Tech http://www.extremetech.com/article2/0,3973,390917,00.asp

Stam, Nick. Inside PCI Express Published by Extreme Tech http://www.extremetech.com/article2/0,3973,522346,00.asp

Teece, D.J. “Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing, and Public Policy.” Research Policy 15 (1986).

Utterback, James M. Mastering the Dynamics of Innovation. Boston, MA: Harvard Business School Press, 1994.

Utterback, J. and Suarez, F. “Innovation, competition, and industry structure.” Research Policy 22 (1993).

Von Neumann, John. “Probabilistic Logic and the Synthesis of Reliable Organisms from Unreliable Components.” In Automata Studies, edited by C.E. Shannon and J. McCarthy. Princeton: Princeton University Press, 1956.