

Building Digital Trust in Today's Pervasive Computing Environment

By Valerie J. McNevin, LL.M., Dr. Bill Worley, Frank Ricotta¹

Introduction

Intentionally or not Basel II is forcing financial institutions to make operational risk a high priority worldwide. Operational risk as defined by Basel for purposes of quantification and capital allocation is “the risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events.” Given the reality of today's environment this risk of loss will be directly or indirectly related to IT systems and processes and the people who use, operate, or manage them.

IT has changed financial services dramatically over the last decade. First, back-end data-processing operations were automated. Then, technology was used to restructure financial store-fronts. At the same time the Internet was used to escalate disintermediation in all financial service arenas. Today, the Internet enables a bank to expand its presence globally without undergoing a requisite expansion of human or capital investment. Now, the movement of money is instantaneous via wireless technologies so that the merger of finance and telecommunications appears seamless to the consumer.

Advances in technology continue to drive financial service changes in four primary ways. First, it enables financial services to exert and maintain a global “presence” with little capital investment. Second, it forces the dynamics and focus of business processes to change quickly and significantly. To serve a global customer base and meet its demand for instant service, business must communicate effectively and efficiently, and it must shift its process so that the customer absorbs as much of the administrative responsibilities, costs, and risks as possible. Third, in order to capture and build on back-office efficiencies, the products, services, and legal relationships of the business must be standardized. To process information more efficiently the business becomes more dependent on 24/7 access to sophisticated information supply chains. Again, these changes reinforce the need to shift risks, costs, and responsibilities to the consumer. As a consequence, it forces business to open the circle of confidentiality to more people as service providers require access to private or confidential information in order to process it. Finally, it redefines market efficiency. From the adoption of real time gross settlement systems and dematerialized securities to the use of straight through processing, XML to T + O trading timeframe, financial service players now literally compete for nanoseconds.

These business process and access improvements coupled with the creation of a global information gathering/processing industry are enabling more people to access financial services, to access these services more conveniently worldwide, to access them faster,

¹ I wish to thank Dr. Bill Worley and Frank Ricotta for the time and effort they have so generously given to this project. This project began with the intent of blending theory with practical application and turning it into reality. Having seen the creative and destructive capabilities of the access that open networked architecture provides, these men have worked to forge a means by which measurable, dynamic trust can be an elegant, affordable, performance attribute of any system.

and to obtain results more quickly. However, access is a two way street. While these technologies give more people access to financial services, this same ability also enables more people to commit financial crimes by stealing, manipulating, destroying, or corrupting information transmitted or stored in these systems. Further the financial services sector fails to appreciate that although these technologies initially lower transaction costs and increase process efficiencies, they magnify system vulnerabilities and significantly increase overall operational risk.

Directly or implicitly efforts to measure operational risk defined by Basel II arguably will impact the “risk culture” of the vast majority of financial institutions around the world. Basel II in concert with the mandate to resolve the corporate governance issues arising from failures such as Enron and the realization that financial services are a prime critical infrastructure target for organized crime and terrorists are driving changes in how the industry governs its “risk culture” from the top down. At the same time legislative responses to Enron, such as Sarbanes-Oxley, are demanding financial entities to change how they “process” risk from the bottom up through better internal controls. Concomitantly, the wholesale adoption of information systems and wireless technologies, especially personal mobile systems by countries and the move to a global 24/7 online market by business, are increasing the scope and breadth of the sector’s external operating risk. The paradox is that while certain advances in technology are enabling us to better understand and mitigate risk, at the same time they making it more lethal and more difficult to control.

One of the most difficult issues encountered in the digital world is *trust*. How can you “know” that the person or entity you are dealing with is in fact that person or entity? How do you know who or what can be trusted? Issues concerning identity, safety, confidentiality, privacy, and integrity consistently haunt every computing space. Based on one bank’s set of online Terms and Conditions (T’s & C’s), this paper introduces a means by which trust-based digital relationships can be built and maintained in today’s pervasive, distributed computing environment. Starting with an overview of the challenges a business faces when migrating to a pervasive computing or networked environment, it identifies and addresses the fundamental components that must exist in a system to achieve digital dynamic trust-based relationships. It then explains how trust-based relationships can operate in vertical “security” space, as well as horizontally in the visual, applications and system spaces. Third, it proffers a new perspective on dynamic trust by identifying the basic attributes that must exist to obtain inherent or pervasive trust² in a shared cognition space and by providing appropriate mechanisms for

² This paper focuses on the concept of trust based systems. The following terms are used to describe the attributes and characteristics of digital trust. These are root trust, pervasive trust, distributed trust, extensible trust and dynamic trust. Root trust refers to the degree of assurance one has in the state of the system each time the software is loaded. Pervasive trust refers to a system state where all network elements provide assurance that they are working with integrity and have not been compromised. Distributed trust refers to occasional computing elements be they mobile, sensors, etc that individually exhibit measurable levels of integrity. Extensible trust is a system where root trust is extended operationally to distributed or non-resident system elements. Dynamic trust as opposed to static trust is an attribute of the system. Dynamic refers to real-time, flexible intelligent trust capabilities that continuously interacts with the system components.

coordinated responses when danger is perceived or unauthorized changes occur. Finally, it describes the means by which a system can be redirected to operate as an intelligent, self-diagnosing and self-modifying predictive environment that enables the system to protect itself in ways similar to how the body's immune system operates. In conclusion, the paper examines the ways that business and performance can benefit using dynamic trust-based pervasive computing.

II. Primary Challenges of Establishing Digital Trust in a Pervasive Computing Environment

Challenges exist primarily in the legal, technical, and operational/management aspects of today's basic e-finance operating framework. In the painting entitled *Wanderer Above The Sea of Fog* by Casper David Friedrich, the artist portrays a man gazing across a sea of fog, uncertain of the future, attempting to peer into the clouds below. This painting poignantly captures today's global market as it ventures into the vague unknown of the virtual world. In essence, the market is poised at the edge of a great cloud of uncertainty, trying to forge new tools that will enable it to navigate through the haze as it transitions ever deeper into the virtual-centric world.

In just a few years the Internet has grown from an experiment in surviving chaos to becoming the most significant delivery channel of goods and services on earth. Whether one is purchasing software, downloading music, buying airline tickets, factory parts, or stocks, it is the channel of choice, particularly for thirty year olds and younger.

Purchasing via the Internet is popular because it is more efficient and less costly to the manufacturer as well as the consumer. To engage in trade over the Internet typically involves the consumer entering into a "point and click" agreement with the merchant. For example, to buy software, the consumer clicks onto the website and then clicks to the purchasing form where he/she inputs his/her credit/debit card information or payment instructions. Once payment is approved, the consumer is directed to the licensing/registration page. Here, the consumer is shown a digital read only copy of the Terms and Conditions (T's & C's). He must agree to abide by these before moving forward in the transaction. Simply clicking on this button constitutes evidence that the consumer agrees to be bound by the T's & C's. Once acceptance is registered, the consumer is permitted to download the product. If the consumer refuses, he is denied access and the transaction ends.

The process is used in forming a contractual relationship online with a financial institution. For discussion purposes the following contains a portion of the substantive provisions of one financial institution's online click and point agreement. It offers a full menu of activities: stocks, loans, bill paying, and online banking among others. To access or use the online channel, a customer must agree to the following T's & C's.

- The bank can change any information on its website anytime at the bank's discretion. Use of or access to the website signifies the customer accepts any changes the bank makes to the website as changes to the original agreement without need for notification or written consent.

- The bank gives no express or implied warranties, representations or endorsements of any nature about anything on its website or linked to its website. It expressly disclaims liabilities for errors and omissions in such material, information, or functions.
- The bank gives no warranties or representations regarding access (in whole or in part) to the website, its contents, its availability, its completeness, accuracy, how current the information posted on it is or that the content is free from any errors.
- The bank refuses responsibility for the upkeep of its website -- from correcting an identified content defect to assuring that it is not used to transmit viruses, worms or anything that could contaminate, damage or destroy any computer linked to its website. It also refuses to represent any level or quality of service such that the customer or the bank may experience transmittal delays, failures, errors or information loss. In particular it refuses any responsibility if a customer experiences loss to his/her investment profits or savings.
- The customer is responsible for evaluating the adequacy, completeness, current status, and usefulness of all services, content, advice, opinions, and other information obtained or accessed on the website.
- The website may not be modified, copied, distributed, broadcast, displayed etc. in whole or in part without the express prior written consent of the bank.
- The bank is not liable to the consumer for any loss or damage of any type from use or inability to use the website or from use or inability to use the system.
- The customer irrevocably agrees to completely indemnify the bank, including legal fees incurred in connection with or arising from the use/misuse of the website and services, breach of the terms and conditions. or any intellectual property right or proprietary right infringement claim made by a third party against the bank in connection with the customer's access or use.
- The bank has the right to terminate the customer's use or access for any reason at any time without notice.
- The customer agrees to be bound by the bank's choice of governing law and that law applies to all aspects of the relationship.

The example above demonstrates how dramatically the Internet is altering traditional concepts of contract law. Historically, certain underlying requirements had to be satisfied in order to achieve fair dealings in forming contracts. At a minimum the terms and conditions had to demonstrate the parties negotiated on equal footing without coercion or adhesion. In addition, the contract had to provide representations and warranties depending on the subject matter of the contract or the level of expertise of the parties.

Here, the bank sets forth a contract of adhesion; it gives the consumer no choice regarding any of the T's & C's, including no choice as to the governing law, while demanding that the customer fully indemnify the bank. The parties do not negotiate, there is no issue of consent, no compensation, no mutuality, and no equal bargaining power evident. The bank assumes no risk, no liability, no responsibility or accountability for maintaining its website, its IT systems, or for damage sustained to the customer's property or information. In fact the bank denies all responsibility, transfers all risk, and disowns any accountability for its website and related links or activities. Moreover, if the

customer wishes to use any of the bank's online services, he/she does so at his/her own peril without recourse to the bank.

The questions that must be asked concerning online or mobile transaction channels are those of first principles. Is not a bank held to a higher standard of care by nature of its fiduciary duties to its customers? Is not a bank given special status, because it holds a regulated monopoly on the public trust? Is not a bank responsible for the safekeeping of a customer's effects, no matter the nature of the goods? Does not a bank carry fiduciary obligations toward the customer's property to conduct its business in a safe and sound manner? Do any of these obligations, responsibilities, or duties change because the bank chooses to conduct business online? If banks choose to use advanced technologies to service their customers in order to be competitive, should they not do so in a way which is fair, promotes the sharing of risks, and protects its customers and affiliates from unnecessary or excessive risk or harm? The over-riding questions then are do not banks, more than any other business, trade on the perceived quality or value of their public trust, and should this trust expectation be lower when transmitting data via the Internet or any other open network architecture?

As contracting in cyberspace evolves it will naturally take on characteristics reflective of this new environment. Such changes are expected. However, these changes should be accepted only when they are fairly and consistently implemented. Historically when the market oversteps its boundaries regulators are forced to intervene to restore balance to the contracting parties. The example here is one of clear abuse. The first question is whether technology can be used to help restore balance? If so, will the market choose to use it for this purpose or will it refuse, and in doing so position itself for regulatory intervention. One thing is clear, the evolution, promulgation, and adoption of trusted cyber systems is essential to restore a much needed balance.

The demand for online-all-the-time availability coupled with a lack of true understanding as to what is necessary to secure these systems is creating significant stress in the industry. One stress point is the question of how much risk the business is facing and how much security it needs to mitigate its risk to an acceptable point. As applications have evolved over the past few years enabling on-line operations, there has been a concomitant extension of technologies, growing from the realization that significant processing must be interposed between today's operating systems, application servers, and the Internet in order to assure trustworthy operations. Such processing includes SSL, offloading, firewalls, DDoS protections, worms and virus filtering, and intrusion signature detection.

Network information systems were designed to move information more quickly and cheaply from point to point. They were not designed to move information safely. In fact when Bill Gates envisioned Windows on every desktop he never imagined that the desktops would be networked. Then along came the Internet. People wanted the desktop to reach out and touch the Internet. This forced Microsoft to insert public domain software into Windows to provide availability and interoperability, which in turn exposes any connected operating system to the same vulnerabilities at the same time. Because network computing is only a little over a decade old, security solutions for these networks are still in their infancy. Given the present design weaknesses and architectural

vulnerabilities of these systems, they are very difficult to secure. Moreover, until now the industry has proclaimed that the cost to redesign these systems is prohibitive. Thus institutions are forced to use band-aid approaches in hopes of stopping system leaks.

- Weaknesses in security approaches include over-reliance on silver bullets. In the 1990s, there were virus scanners and firewalls, then PIN passwords, SSL, and PKI. Each of these technologies has pros and cons. Not one in and of itself is a complete security solution. Each can provide a measure of protection. Yet how strong that measure is depends on the skill of the person attempting to access the system.
- System threats exist on three levels: software, hardware, and people. Software threats include malware, viruses, worms, Trojan horses, and bots. Hardware problems can result in denials of service, distributed denials of service, crashing or compromising the hardware. People remain the most significant problem and include insiders (those with some level of approved access) and outsiders (those without approved access but who manage to access the system through illegal means).
- Cyber crime typically falls into the following categories: extortion, identity theft, stock market manipulation, credit card theft, check theft, and salami slice. At present approximately 39 million people in the US alone are victims of identity theft. Identity theft was first recognized by the law in 1996. Since then it has become the fastest growing crime in North America and Western Europe. Wherever advanced technologies are widely implemented, identity theft quickly follows.

Given the risks involved, the question is whether the market is ready for high value, high volume activity, or whether as a matter of fiduciary responsibility, public trust, safety, and soundness, it is better to use the technology for certain activities until risks can be contained and better security solutions are available? In the final analysis, the issue is whether we can manufacture trust and substitute it for public, social, and commercial trust?

From a technical perspective, the over-riding challenge to establishing digital trust relationships in today's pervasive computing environment is the inability or failure to establish root trust in the operating system, and to extend chains of trust into all critical applications. Root trust means there is an originating point of control and authentication that provides significant assurance of confidence, integrity, and reliability. Today's operating systems are not designed to verify from the point of origin, otherwise known as "booting up" the system. In other words it cannot determine that the system is in fact what it claims to be. To boot up securely requires all system code to be verified and authenticated as it is downloaded. Basically the operating system undergoes a virtual physical exam to ensure that nothing has tampered with its DNA. After the verification process is completed, a similar verification and authentication must be performed on all applications that will be involved in interactive transactions. Only then should the system

begin to determine whether the party initiating the communication is the person he/she claims to be.

Root trust combined with ubiquitous, uninterrupted chains of trust is critical to creating and maintaining value from a legal and operational perspective. If the point of origin cannot be demonstrated as authentic, the trust value chain is at risk from the start. If the trust does not extend into each application, risks of incorrect transaction execution cannot be eliminated. Without pervasive trust, any communication is potentially suspect as to the identity of the sender, the integrity of the message content, and the actions actually performed for a transaction.

Without pervasive trust, the legal and technical challenges can result in an operational nightmare. To understand the importance of the trust chain value, an entity's processes must be clearly captured and each element of its process identified. Process defines the value in the chain. Process integrity and real-time control optimizes the chain's value. Understanding transaction and process flow is essential to establishing and protecting the trust chain. Every element from and including the point of origination must be identified, monitored, tested, and verified to ensure message integrity and point of origin authentication. At any point in the chain, if an entity gains access to the system that access must be identified, verified, and authenticated in real-time. If this fails to occur or if unauthorized access is identified, the value chain is detrimentally impacted and the risk of loss, manipulation, or control is suspect.

Once the value chain is understood, protecting the trust chain becomes one of the most critical data governance issues facing a company. Fundamentally, data or information assurance is a process and communication issue. How well it should be accomplished translates into the company's performance standard.

As a value proposition, in today's business model where customers interface with a bank's systems via open network architecture far more than via the public switched and closed networks, the bank needs to critically assess all access points and determine the level of trust it is willing to deposit in each entity accessing its system. This should be based in part on the bank's knowledge of that system as well as the value of the data traveling between these points.

Manufacturers add to the confusion by insisting their products are impenetrable or impervious to attack. After significant investment of time and resources, the purchaser determines that the technology cannot perform as represented. Unfortunately, this often occurs because 1) there is a disconnect between the design team and the sales team, meaning that the sales team really does not understand the technology, and/or 2) there is a serious difference between advertised product capabilities and the user's functional needs. These problems are shared by every industry, and though they are acknowledged, they are beyond the scope of this paper.

Another way of viewing the problem from an operational or personal perspective is to think of an entity's risk culture as the mirror image of its trust level. If there is a low level

of perceived trust in the organization, then the entity has a high risk environment and is adverse to accepting additional risk. On the other hand, if it is operating at a high trust level, it can accept risk more easily. The goal is to allocate risk where it can produce a positive effect for the company. The way to do this is to create a high trust operating environment so the company can transfer its risk load from the operating channel to a profit making one.

In addition to the inability of today's operating systems to provide pervasive trust, there are three design challenges that prevent a turnabout without radically new thinking. The first design challenge involves the extensive vulnerabilities inherent in the design of operating systems. This concept is elaborated on in the paper on Time Value Based Trust. The most direct method of compromising a network system is to attack its operating system. This provides near-instant reaction time and very predictable results. To make matters worse, the same problem is likely to exist in all computer systems of the same type, delivering results that are nearly universal in nature. Network protocol(s) design is a second design flaw. Most network protocols are designed to communicate promiscuously with other computers without verifying their trustworthiness. Third is forced trust violations. In forced trust violations an entity leverages a trusted vehicle to access a system then abuses the entity's access privilege to enter unauthorized areas of the system. Violations occur by obtaining "secrets" from trusted insiders by stealing encryptions, passwords, through eavesdropping, human interaction and information exchange, data mining, sabotage, corporate sabotage, internal sabotage, and extortion, to name a few.

Today's business model operates to exacerbate the problems created by these design flaws. Corporate assets are transmitted and stored in virtual data systems either under direct control of the company or outsourced to an environment over which the company has no hands-on control. Two critical issues emerge from this situation. First, how can the outsourcing entity be assured that its provider is properly protecting the data? Second, and perhaps more daunting, is the issue of how can the outsourcing entity assure itself in the face of conflicting performance requirements for different clients that its outsource provider is meeting its expectations? As a result the outsource provider is another means by which a forced trust violation can occur. Because of the nature of open network architectures wherever a system's protection is at risk this is known as the weakest link. This means the degree to which you trust an entity, and evidence this trust by allowing your systems to communicate with its systems without protective processes you expose your system to being compromised.

As commercial operations migrate from legacy systems to enterprise networks, and on to occasionally connected (OCC) or pervasive computing environments, the dilemma only increases in magnitude. OCC is a new form of networking that refers to the ability of a computing device to travel with complete network transparency. "Pervasive computing is a term that describes numerous, casually accessible, often invisible computing devices, frequently mobile or embedded in the environment, connected to an increasingly ubiquitous network infrastructure composed of a wire core and wireless edges." (NIST) In other words, pervasive computing is communication among a varying number of

individual computerized nodes that share knowledge to produce correlated results. It is significant in that it does not rely necessarily on any type of central processing and is incredibly powerful in its corroborative reporting and communications capability. Pervasive computing implies that process control is distributed to a system's elements wherever these may be housed. However, because it has no predictable boundaries, the opportunities to breach or access a system without authorization expands exponentially.

As the global network develops, its ability to establish extensible, distributed trust is critical to establishing commercial integrity, trust, and confidence in the network. The question is whether technology can provide an entity the means to monitor and control system elements not under its direct control real time so it can measure the degree of trust in any transaction? The answer is yes. This will be explained further below.

Having failed to manufacture legitimate trust in the present environment, the industry appears to be without the ability to establish extensible distributed trust in the pervasive computing arena. Nevertheless, to do so is a fundamental prerequisite for the next stage of Internet development.

III. How Digital Trust Operates

In order to create a trusted computing environment it is essential to understand how trust as a process or a technology operates in the non-virtual realm. For purposes of this discussion we will look at social, commercial, and public trust to determine if trust can be manufactured and replicated through technology.

Trust is a state of being. It exists as a matter of degree. In human life it requires a good deal of energy and effort to build trust and relatively little to abuse it. Once abused it may take a long time to rebuild. Failure to maintain trust relationships can lead to lost productivity, degraded performance, litigation, and other costs. In looking at the relationship between trust and growth, Zak and Knack (2001) shows that trust depends on the social, economic, and institutional environment in which transactions occur. In effect they claim it is a derivative or outcome of transparency, accountability, and responsibility. Trust operates similarly in a computing environment. It takes time to identify and establish a trust environment and a trust process.

Trust is an established benefit. First, it is a performance enhancer. Numerous studies conducted over the years establish that where high levels of trust exist, commercial activity flourishes. Conversely, if the trust level is low, commerce fails to develop well or consistently. As important, distrust creates commercial liabilities. To date, despite efforts to manufacture trust, the Internet and online markets continue to be a trust-challenged commercial environment. Once this trust is established, online commercial activity will experience a more consistent growth pattern. In the meantime establishing legitimate digital trust mechanisms still claims to be the holy grail of e-commerce.

Example: Computer A and Computer B are separated by one thousand miles. A initiates dialogue with B. B is an access point to an online bill paying service. A is the customer.

A enters a payment order to B to pay a certain bill. B needs to be assured that A is authenticated and authorized to enter the order.

Digital trust mechanisms must work well in a dynamic, interactive environment. At the same time they must be consistent, predictable, and flexible. So in the example above when B verifies that the message was originally sent by an authorized party via A, and that the message exhibits inherent trustworthiness or integrity (i.e. it is the message that was originally sent without compromise and the party who sent it is in fact the party it claims to be), B is authorized then to proceed with the next steps to execute the payment order for A.

To be effective trust must be able to operate between B and Computer C at the bank where the funds are held for A. The same process occurs. Verification, authorization, and authentication are elements of Trust that the users are assured exist by measurable, demonstrable means.

What is trust and why is it important? Trust is a vital part of the holistic risk management system inherent to humans. Trust transforms. It is the lever of which Aristides spoke. Collective trust significantly transforms negatives into positives. It begins as a personal choice, builds into a lifestyle, and eventually translates into a public system of beliefs. Living trust generates its own economics. It creates a depository of will. Trust among people grows into social capital and creates wealth. As trust grows, responsibility, accountability, and transparency also grows and completes the virtuous circle. It can be categorized in many ways: personal, private, public, commercial, social, or implied, to name a few.

We are all familiar with social trust. Social trust is a risk management tool that relies on information gathered from inter-personal and social contacts. It provides an assessment of whether the other party is committed to acting in good faith. It permits decisions on whether or not to accept the risk that the other party will deliver and, if possible, takes actions to minimize risk if expectations are not met. Today, one may choose to do business with an entity, though it may cost a little more, because over time, one has developed a relationship with that entity and feels safer by staying with it.

Commercial trust operates in much the same way as social trust in respect to due diligence yet results from the need to expand opportunities and explore new markets. It substitutes man-made tools and techniques to investigate and perform due diligence on the market. The information gathered sheds light on the investment environment. Using these tools, it is possible to measure the market's strengths and weaknesses and form an opinion about the extent to which one believes the market shares one's values and principles. The ability to accept an investment risk is in part colored by the ability to rely on the market's integrity and credibility.

Three tools and two techniques are heavily relied on by international and institutional investors to provide a framework within which they manage their commercial trust relationships. These tools are: (i) the laws of a country; (ii) the rules and regulations that interpret those laws; and (iii) the processes used to resolve disputes and business failure. The techniques measure performance. One measures the degree to which regulators

require the market to comply with the laws and the degree to which business abides by the law. The second technique measures how the market and the regulators handle expectation failure.

In designing risk mitigation instruments or environments five important questions should be asked.

- Can we identify the risks?
- Can we trust our counterparts?
- How do we trust them?
- How can trust be represented in the transaction?
- How can trust be enforced?

Written contracts are one of the oldest instruments evidencing the transition from social trust to commercial trust. Written contracts were introduced into society to provide a means by which parties could record their expectations and set out their remedies if there was a failure to perform. It enabled parties who did not know each other to take the contract to a third party, be it clergy or judge, to resolve disputes and obtain some type of satisfaction in the event of expectation failure.

Today most financial systems operate at least to some degree through electronic contracts that are managed and monitored by automated systems. We are missing a technology that enables us to develop virtual or manufactured trust.

IV. Characteristics of Dynamic or Real-Time Trust

Establishing trust throughout a complex system comprised of data base servers, data and transaction archive servers, application servers, Web servers, Web Edge appliances, as well as stationary and mobile clients is a significant challenge. However it is a challenge that can be met by deploying the emerging *Secure by Design* system components and by systematically applying the *Design for Dynamic Trust* principles to the remaining components.

As major operating systems and system components have evolved, their structure has been guided primarily by goals of generality, portability (to multiple hardware platforms), and continuous availability, rather than goals of security and pervasive trust. Thus, complex systems built in a straightforward manner from these system components exhibit continuing vulnerabilities to network-based attacks. Maintenance of such system components also requires careful planning, testing, and resources simply to keep up with and apply what has become a continuing stream of software security patches while maintaining reliable operations. Despite all these patches, security penetrations persist with alarming frequency.

One important realization about system components based on major general-purpose operating systems is that they cannot provide root trust. Essentially their structures are still based on the same two-level hardware privilege model introduced 45 years ago when IBM introduced the System/360. Per this structure critical system components execute in a “privileged” mode. In this mode the software is able to exercise complete control over

the hardware. Application programs however execute in a “non-privileged” mode that restricts hardware access only to those facilities allocated to the application by the operating system.

When software executes in privileged mode in these systems, the system has no defense to ward off an attacker’s insertion of intentional malice into the code. Indeed, an attacker’s primary objective is to introduce code that incorporates the attacker’s malicious intent into a system so that it executes in privileged mode. But if a system can provide root trust, it boots to a precisely known state, and is immunized from the dynamic introduction of any additional executable software.

Over time, the number of system components executing in privileged mode has grown far beyond any expectations of the original designers of the System/360 protection structure. The resulting complexity of the software components and of the transitions to/from privileged mode has reached the point where it is no longer comprehensible. In turn this has led to an unquantifiable number of system vulnerabilities and avenues of attack. In time one reasonably should expect advances in the security properties of the major systems. But for now, other system components are required to establish one or multiple points of root trust. And, one or more points of root trust must exist in a system designed to reduce quantifiable risk.

Current publications that discuss emerging “Trusted” computer systems are called “software closed”.³ Such systems prohibit the introduction of any additional executable software. A system component capable of providing root trust also can be designed to be “software closed”. Further, such a system component can employ a simple, special-purpose, minimal operating system, designed specifically to scan and move information quickly and safely, and to use the most advanced hardware features to eliminate all vulnerabilities and avenues of network attack. Such a system component which can process high value information while assuring root trust and pervasive trust, can be considered to be *Secure by Design*.⁴ Equally important, such a system component can be a performance enhancer at the operations, management, and product levels because high overheads associated with general-purpose operating systems are eliminated.

One or more such systems can be configured at strategic points within an enterprise environment. These system components act as anchor points that distribute trust to the surrounding system components. The anchor points provide the foci for centralized command and control of the system’s elements. They authenticate identities and configurations of other general system components; authenticate identity and authorization credentials of those who administer or otherwise interact with the system; secure transmission of information between regions of the system; continuously scan transaction requests for malicious content; identify policy violations and unauthorized requests; and coordinate continuous monitoring of the integrity and state of the client and server systems.

³ IEEE Security & Privacy, March/April 2005.

⁴ Secure by Design means that the system is able, by design, rather than through an add on feature, to provide demonstrable, measurable trust, be it root &/or distributed trust.

The principles that govern configuration and management of clients and servers include best practices for defense in depth and layered security.⁵ All unused system software components must remain disabled, and all used system software components must be set to execute with minimum privileges and access.

In addition to these best practices, the governing principles for the system structure being proposed mandate that agents be installed and refreshed dynamically on every participating client and server. These authenticated agents continuously monitor the configuration, health, and data flowing into and out of each particular system element. Technology now exists which can dynamically monitor strongly authenticated agents in specific system elements. Once installed, such agents can continuously monitor system configuration attributes and behavior.

Initially, monitoring results can be communicated to root trust anchor points. In effect, the behavior of all system elements that are not *Secure by Design* can be *Designed for Dynamic Trust*⁶ by being continuously monitored by authenticated agents in communication with root trust anchor points. Over time, such monitoring agents can evolve peer-to-peer trust relationships among themselves and coordinate control, error identification, and recovery operations in a far more fine-grained manner.

A design of this type of system employs numerous security technologies. Particularly attractive is zero knowledge authentication. At its most fundamental level, zero knowledge authentication ensures that a node can be trusted and that the person requesting access is authorized to ask for and receive the information requested. It differs from other authentication technologies for two reasons. First it can prove a person's identity through possession of a unique secret without revealing information about the secret used to verify the identity. Second it is able to create a unique identity that belongs to the user and the system component being used. The fundamental core of the system is its ability to create a secret, generated through a combination of physical and logical characteristics of the system and the person attempting to use the system. Zero knowledge authentication has the further advantage of obviating the complexities of a pervasive PKI framework.

All communications between agents and root trust elements must be protected for confidentiality and integrity by strong cryptographic methods. Pervasive authentication empowers the system to continuously monitor and report in real-time if there is incorrect or unauthorized behavior by the system or its users. It also can detect malfunctions and initiate restorations of the agents.

⁵ T. Glaessner, T Kellerman, and V. McNevin, *Electronic Safety and Soundness – Securing Finance in a New Age*, World Bank Paper No. 26, 2004.

⁶ Designed for Dynamic Trust means that a system which has not incorporated security inherently into its design can be augmented to provide a degree of dynamic trust with the appropriate changes. Because it has design flaws it cannot provide the secure by design trust but can perform better than a system without any augmentation.

Today's security solutions fall short of protecting the enterprise because they still focus on securing data centers and the employee's access to this center through an "own and control" approach that can only protect the company from third parties at its perimeter. As a result, they are not keeping pace with the needs of the enterprise. In today's environment, the communications backbone of any business increasingly includes mobile devices. These are deployed to accompany the employee wherever he goes. As a result, these assets cannot be controlled physically by the company, nor can they be protected by perimeter solutions. In addition, enterprises face increasing threats from the "trusted insider". Today this includes employees with significant IT skills.

On the other hand a dynamic root, distributed, and pervasive trust framework provides trusted access, collaboration, and responses across any enterprise, including fixed and mobile assets. Second, it ensures that an asset can be trusted and the information being sent has not been tampered with in any way. Third, it aids in creating multi-levels of information access by ensuring that the user is authorized to ask for and receive the information requested. By providing intimate knowledge and information about the system at a very granular level, it prevents system spoofing and trusted insider attacks as well as attacks from third parties that have successfully penetrated today's security perimeter solutions. Fourth, distributed peer-to-peer agents are able to dramatically form and reform in response to specific threats and tasking.

Extensible distributed trust means the system is measurably known to be free from attack, compromise, or manipulation at the time the data was created, and can prove the data made the journey from origination to point of control at the edge without loss of integrity in a measurable way. Finally, extensible distributed trust means that those elements of a particular transaction flow, even though they are not under the ownership or control of the company, can be monitored and protected from attack or unauthorized access.

Today, an important window of opportunity is open to participate in the formation of the global legal and ethical architecture of cyberspace, and to infuse this structure with ethics and values thereby helping to provide it with a moral compass. This input is critical as we navigate our way through the cloud of uncertainty. The system structure articulated earlier in this paper can assist this formation by defining minimum standards to be provided by a qualifying cyber system. Such standards can reduce the uncertainties attendant to formulating needed ethics and values.

IV. Benefits of Mitigating Risk

The best risk mitigation method to employ is trust building. Defining risk within and by the context of a particular venture is essential to being able to design mitigation techniques to better control it. Second, it is essential to understand the roles being played within the theatre of the contractual agreement. Identifying and clarifying the obligations, roles, responsibilities, and interdependencies of the various parties is essential. Similar identification and clarification is required for elements of the cyber systems.

It is critical to establish who takes on risk and for what purpose. One necessary exercise in any transaction is to evaluate it in terms of the parties, their roles and responsibilities,

their assumptions about possible outcomes of the transaction, and how the risk will be mitigated. Determining whether the risk is excessive, and identifying what can be done to mitigate the level of risk to which a party is exposed is crucial to whether ethical trade will thrive in this environment. The law must provide guidance as to what is right and what is inappropriate.

1. What risks are involved in this transaction or contract?
2. What are the roles of the various parties and their responsibilities?
3. What are the roles of the various elements of the cyber systems?
4. Are these risks excessive? How is excessive defined for this transaction?
5. How can the risks be mitigated and or eliminated so that they are tolerable?
6. What does risk mean to the various parties?
7. What are the roles of the various parties? Specifically, how does risk affect their role in the transaction(s)?
8. What is the compensation for the risk undertaken?
9. Can the risk be transferred? Can the responsibility be transferred?
10. What role does information play in mitigating or managing the perceived vs. actual inherent risk in a particular transaction?

Once this information is determined and the legal and other constraints within which the agreement must operate are identified, the transaction can be designed or structured to mitigate and even potentially eliminate much risk. The most widely adopted means used to effect a transfer of risk or responsibility is to share it; i.e., enlarge the circle of risk participants or decrease the level of risk, and establish trust.

Historically, contractual equations between parties resulted from a negotiated effort to balance risk and responsibility so that the parties could obtain the particular benefit of the transaction without incurring unnecessary risk or unexpected responsibility, which in turn might lead to diminished profits and added costs. Today, much of commercial litigation is built on the concept of risk and how it was not properly handled during the design of a transaction.

In the earlier example a few small changes can turn a negative instrument into a positive agreement that would foster a supportive, trust-based relationship. First, if the bank is so concerned about using a virtual environment, why is it being used? If it raises the risk profile that much why is the bank exposing itself, its shareholders, or its customers?

The bank could maintain an information only website explaining why it is not opening a website storefront. The bank could advise its customers that use of the website is discouraged unless the client has access to a secure line. The bank could use the opportunity to raise consumer awareness and education about the dangers of the Internet. Further, the bank could provide security software and require the customer to install it on his system before he is allowed to initiate activity with the bank's website. By so doing, the bank mitigates its risks and provides a public service for the entire Internet community. Furthermore, if a bank chooses to go online, it should be responsible for

maintaining its website and ensuring all information on the website is accurate, timely, and complete. It should be prepared to inspect and correct any defects on its website once notified about such. Banks could even provide an incentive—for customers that do find errors or problems and inform the bank. This approach recognizes and rewards virtual citizenship.

Building trust suggests that the parties agree to engage at certain levels of risk at certain stages of the relationships or agreement. Ways to build trust include providing the following. This is not meant to be an exhaustive list. At a minimum trust implies the responsibility to communicate and agree on the following:

- Defined terms, concepts, and principles.
- Standards, rules, or principles to be used as benchmarks
- Transparency
- Consistency
- Means and method of resolving disputes
- Choice of Law
- Governing Law

Educating customers and working to raise their awareness of the dangers lurking in the passive use of possibly unsecured technologies is a more time consuming approach; however it is much more beneficial to the commercial environment. This approach keeps the real process focused not just on the relationship but the quality of the relationship as well.

Real-time root and pervasive trust, coupled with dynamic trust capabilities need not be expensive. In fact such trust attributes can increase performance and optimize use of the technology in a cleaner, simpler more capable and responsive computing environment. Given such cost and performance, such a system can provide a safer and sounder environment for the development of low value payment systems and micro payments. By designing a simpler operating system high volume, low value payments can move at a highly accelerated basis with less oversight and better intrinsic knowledge of the system. At the same time it provides the ultimate environment for moving real time gross settlement payments as well as the settlement and clearance of securities.

According to one database of operational risk loss events, the total amount of losses in the financial services industry is in the range of \$400 billion over the past ten years. This amount is based on events in the public domain and is estimated to be perhaps just 10% of the total loss amount. Financial crimes in the UK now are more than 2% of GDP. To understand, much less manage the complicated and erratic nature of this global “always on” market requires serious efforts from the legal community. Risk mitigation is not just a checklist. At its most fundamental, it is a “quality of business environment” or moral imperative.

V. Summary & Conclusion

Even as Basel II and other drivers set expectations that the industry will work to better understand and measure its operational risk, all institutions that are critically dependent

on technology need to undertake this discipline seriously. How an institution designs its risk management strategy inevitably will rise or fall on how well its risk relationships are defined and how well they are covered through legally enforceable mechanisms that are grounded in first principles and fashioned with the desire to create a trustworthy environment through risk sharing rather than to simply transfer risk and responsibility unfairly onto another party.

Globalization widens the scope of opportunities for trade and production. It also exposes us to greater risk. The traditional risks associated with market, credit, political, and legal risk are magnified in a global pervasive computing environment. Technology coupled with globalization exponentially increases risk and opportunity. On the other hand advances in technology are providing us with an opportunity to better reason our way through risk by identifying, anticipating, measuring, and ultimately mitigating it better.

As a state of being, trust too is a technology, an ancient technology that should be revisited and revived into the virtual framework of our collective future. Such trust must be dynamic and flexible. It must start at the root, extend out and distributed throughout all system components. It must create a pervasive state of trust within the system.

We must take time to create the future. In creating the future, we must give place to trust. We must fight to preserve the sanctity of the human spirit and the sacredness of trust in ourselves and each other. We must practice the ethics of trust, for without trust in ourselves, in each other and in something greater than ourselves there is no future.