NeoCatena Networks Inc.

>> Next Generation RFID Security >>

RFID SECURITY SYSTEM

Defeating Attacks against the Digital Supply Chain and other RFID Systems

CONTENTS

RFID TECHNOLOGY TODAY

Departing from the well-established EAN barcode standard, major retailers like Wal-Mart in the US and Metro Group in Europe as well as large pharmaceuticals like Pfizer or Merck are migrating to RFID technology.

However, RFID is not only used as an enabling technology for the digital supply chain - RFID systems are used across industries for many different applications. Other examples are electronic payment and access control systems, RFID implants for livestock and the electronic passport as defined by International Civil Aviation Organization (ICAO). A lot of different standards are emerging, most prominently around the electronic product code EPC, but custom solutions are not uncommon and best practices are not always in place yet.

RFID is an automated data collection technology that uses radio frequency waves to transfer data between a reader and a RFID tag to identify, track or locate the physical item it is attached to ("Internet of Things"). RFID does not require physical contact or line-of-sight between the reader and the tag. RFID tags can carry a significant data payload ranging from a few bytes for low-cost tags to several kilobyte for high-end tags, and many chips not only allow reading but also modifying the data stored on the tag.

RFID SECURITY RISKS

RFID enabled applications have significant benefits over applications based on barcodes, but the technology also introduces a number of new business risks. Flawed designs are not uncommon in today's RIFD systems and best practices and security policies are still emerging. Considering the potential vulnerabilities of backend systems attached to an RFID infrastructure the problem becomes even bigger.

RFID tags are always an integral part of a larger IT system and should be seen in this context. Typically, RFID tags contain a unique ID (for example an EPC product code or any other unique ID). Often this ID can be altered on the tag. In addition, many tags come with user data, a separate memory area that can be used to store additional information directly on the tag, for example, expiration dates for perishable goods, or the dollar value of tickets used in public transport systems. High Frequency (HF) tags traditionally come with sizable user memory, but now we see a similar development with Ultra High Frequency (UHF) tags: Major players in the EPC arena are about to release new versions of their RFID chips containing 512 bits of extra memory.

Given a compatible RFID reader device, anyone can freely read and modify data stored on these RFID tags without the legitimate owner even being aware of it. RFID auditing tools like RFDump (www.rfdump.org, released as free software by the founders of NeoCatena in 2004) can be used to explore the weaknesses of existing RFID infrastructures.

On-Tag Encryption

On-tag encryption, if used at all, is typically proprietary and weak due to physical limitations and cost constraints. Once the encryption features of an RFID tag are broken, the tag becomes nothing more than an ordinary data tag that can easily be manipulated with tools like RFDump. For example, the security features of the Mifare Classic chip, which is used in public transport systems and building access control world-wide today, has recently been compromised. At the Chaos Computer Congress 2007 Karsten Nohl from the University of Virginia presented the results of his research. Nohl had analyzed the Mifare chip layer by layer under an electron microscope and reverse engineered significant parts of its proprietary encryption logic revealing major design flaws showing how easy it is to break the chip's security features. With the dollar amount of the ticket directly stored on the tag, ticketing systems based on this chip, like the Oyster Card in London or the Charlie Card in Boston, are at risk. An attacker could attempt to either clone a ticket or change its value to gain illegal access to the service provided. Similar cloning and tampering scenarios apply to other open loop applications as well, including hotel key cards, ski lift and event tickets, electronic payment systems and the electronic passport.

The Mifare Classic chip is the dinosaur among the RFID crypto chips – it was first released in 1994 and its proprietary crypto design has major flaws and well known weaknesses. In recent years, however, several researchers and companies have started working on so-called "light-weight cryptography" solutions for RFID. The idea is to re-evaluate existing crypto primitives andinvent new algorithms specifically designed to achieve reasonable levels of security for low-cost passive RFID tags. This is a major challenge considering the limited computational resources as well as power and timing constraints on these chips. For example, chips used in today's low-cost EPC tags contain up to 15,000 gates. Only a fraction of these are available to implement crypto functionality, the rest is required to implement the tag's state machine, memory etc.

RFID SECURITY RISKS (continued)

Strong private key crypto systems on the other hand require at least 20,000 - 30,000 gates alone when implemented in hardware. Still there are a number of new cryptosystems available that can operate under these restrictions and their inventors claim security levels comparable to RSA and other well-known strong crypto systems. The jury is still out on whether these claims are true. It remains an open question whether it is possible to deliver strong security with such limited resources at all.

RFID Malware

Storage capability on RFID tags is limited but sufficient to carry attacks commonly known in the IT security world (e.g. SQL injection, buffer overflow, string format attacks etc.). Even RFID based Malware is possible, as demonstrated by Tanenbaum, Rieback et al. in their paper "Is Your Cat Infected with a Computer Virus". Essentially we see the same attack patterns and exploits that have been studied in the network security community for a long time emerging in a new domain.

Protecting Internet communication channels using firewalls has become standard, but in today's typical RFID implementations very little is done to protect the edge of the network. This opens the door for various kinds of attacks using the RFID communication channel as an attack vector. Often RFID systems are perceived as closed loop applications with no exposure to the outside world. However, in a digital supply chain scenario, what happens after items or packages have been labeled at the manufacturing site and before they reach distribution centers or warehouses? While in transit, tags can easily be accessed and then altered or exchanged. Fake products and tags could be injected into the supply chain, a serious concern in the pharmaceutical industry. In addition, the possibility of an inside attack should not be underestimated

Attacks Against RFID Applications

In summary, today's RFID applications are exposed to numerous technical risks with direct consequences for the business such as revenue loss, customer safety, fraud, brand damage, business continuity and liability. In addition, these risks affect legal and regulatory requirements as defined by SOX and Basel-2. These and similar legislation in various countries require the operator of an RFID application to integrate RFID into their risk management process and analyze RFID specific risk scenarios and their effect on business continuity.

Some of the technical risks and their implications are listed here:

- Cloning: Duplicating or manipulating RFID tag data to create identical copies or variations of RFID tags that will be accepted by an RFID application as valid. Goals could be to gain illegal access to a restricted area, inject counterfeit products into a digital supply chain or change price tags at the Point of Sale (Cyber Shop Lifting).
- Code Injection Attack: Manipulating RFID tag data so that it contains malicious code or code fragments (RFID Virus) with the intention to change the course of execution of backend systems or databases processing the RFID data. This Malware attempts to exploit vulnerabilities caused by sloppy software implementations bringing the system into an undefined state, or crash it otherwise (e.g. because of missing error handlers, vulnerability to buffer overflow, string format or other injection attacks). In particular SQL injection attacks may be used to alter any record in a SQL database.

RFID SECURITY RISKS (continued)

The goal could be to gain illegal access to backend systems, or to disable or destroy it entirely.

- Denial-of-Service Attack (DoS): Almost all resources in an RFID system can become target of a DoS attack, including tag, reader, or back end system / edge server. Attacks on the air interface include shielding tags, flooding the reader field with a multitude of tags or selectively jamming the reader field. The goal is usually to sabotage specific resources of an RFID system, such a digital supply chain, effectively making the system unavailable to its intended users.
- Man-in-the-Middle Attack: Impersonating either a legitimate reader or tag to collect RFID data and metadata, or actively manipulate data sent between tag and backend system e.g. to perform code injection attacks, gain illegal access to a restricted area etc.
- Eavesdropping/Scanning on the RF channel using special reader devices and sensitive antennas to collect RFID data, e.g. to collect information about a business (corporate espionage) or individual (profiling, tracking tracing, privacy violations). This type of attack is often used as a helper attack serving as prerequisite to perform further attacks such as a man-in-the-middle attack etc.
- Physical Attack: The simplistic architecture of low-cost RFID makes these types of attacks feasible once again. There are invasive physical attacks (microprobing, focus ion beam editing etc.) and non-invasive attacks (power/timing analysis, radio finger printing etc.). The goal of these types of attack could be to reverse engineer proprietary crypto algorithms and extract keys.

Some of the attack types listed above, the physical attacks for example, are costly and time consuming and require a high degree of technical expertise. These attacks are often impractical and only performed as a proof-of-concept in the research community. Other attacks only make sense in combination with other attacks. An eavesdropping attack, for example, could be performed to learn about the communication protocol between tag and reader to later perform a code injection attack.

Different industries use different RFID technology and face different kinds of business risks. For each vertical industry a careful analysis is required. A good starting point is a detailed threat model identifying possible incidents as well as the probability and cost associated with their occurrence. Risks can then be coupled with recommended countermeasures and solutions providing an adequate level of security can be tailored specific to an industry's or even and individual user's need.

DEFEATING ATTACKS AGAINST THE BACK-END

One of the major attack types against RFID systems are tag data manipulations attempting to perform an attack against backend databases or applications using some form of code injection attack. The reality is that most RFID systems do not account for safeguarding the RFID communication channel at all. As a result these systems cannot spot manipulated RFID tag data and react accordingly.

To protect backend systems from these types of threats an independent authority is required, ideally a dedicated security appliance, restricting access and preventing arbitrary data from being injected into these systems.

The solution should integrate transparently and seamlessly into a given RFID system directly after the RFID reader and in front of the backend (Edge Server or RFID Middleware). This way the solution can inspect tag data before it reaches the backend and can block it if it poses a threat or let it pass through if the tag data is ok. Thus, even if a malicious tag enters the reader field, the tag data will not propagate beyond the security appliance and the backend will never even be aware of its presence. Instead normal operations will continue as if nothing happened, effectively shielding back-end systems including Middleware, databases and ERP solutions etc. from Malware and attacks via manipulated RFID tags.

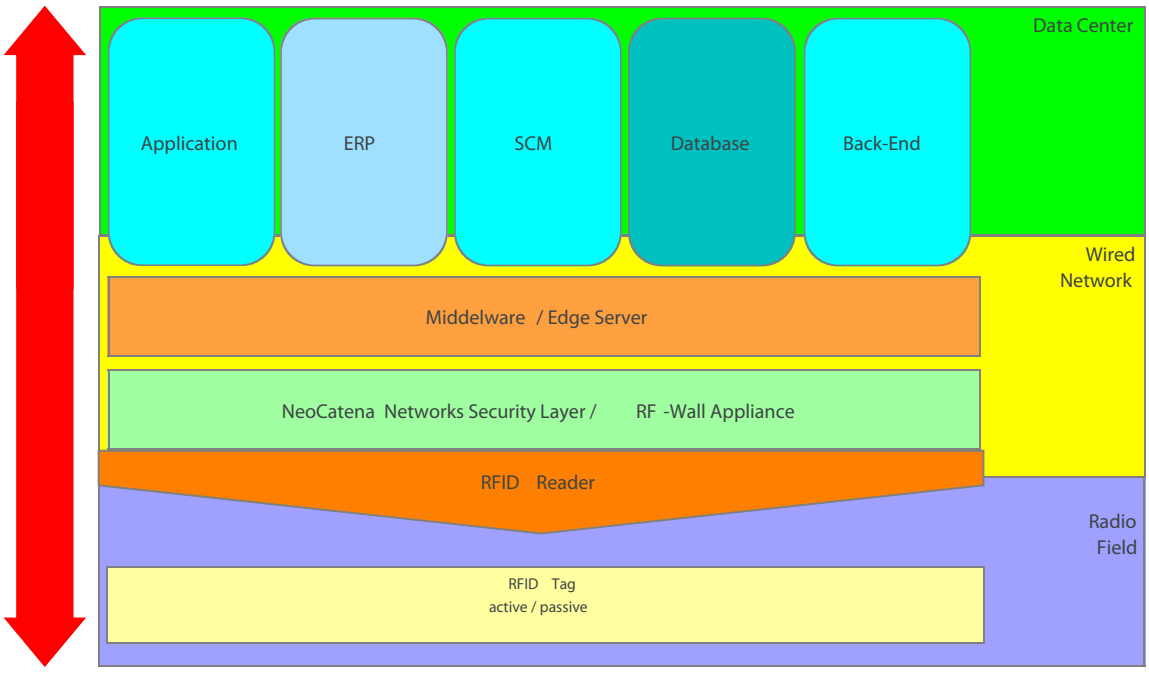

Diagram 1: Logical Concept

DEFEATING ATTACKS AGAINST THE BACK-END (continued)

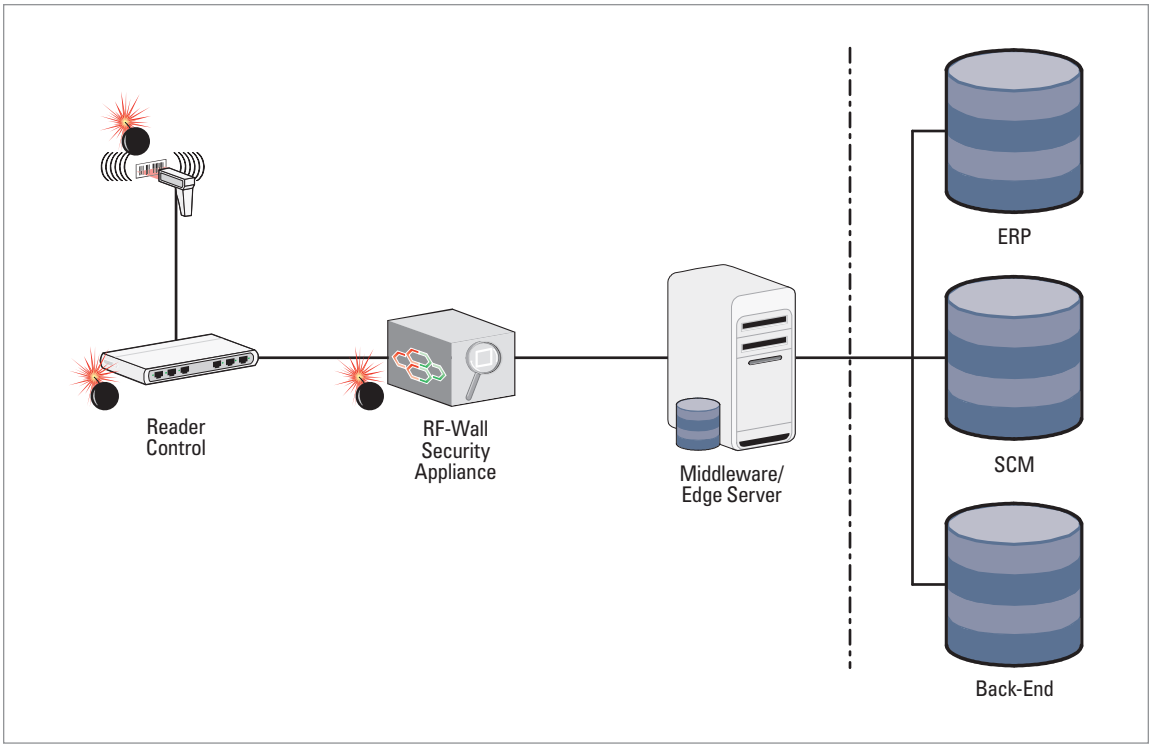

Diagram 2: Protected Backend

The central functionality of a RFID security solution can be compared to a combination of firewall and virus scanner / intrusion detection system (Solution 1). Tag data is analyzed for validity and conformance with defined data formats (positive tests) and potential exploits (negative tests). Only valid tags passing all tests should be handed through to the back-end systems for processing. By default the solution should provide components to validate a number of standard data formats (e.g. valid EPC product code, well-formed XML etc.) as well as generic exploits and known attacks (including SQL injections, buffer overflow attacks, string format attacks etc.). Over time the intrusion detection component should get a sense for normal RFID traffic and hence be able to detect unusual or suspicious traffic patterns on the RFID communication channel. Finally, the solution should allow for regular scheduled updates of its attack signature databases from a central location.

With RF-Wall, NeoCatena provides a solution that meets and exceeds the requirements outlined above. RF-Wall is implemented as an appliance intercepting all data coming from one or more RFID readers before it can reach the middleware or backend that would otherwise be directly connected to the reader. As an added value, the RF-Wall may be utilized for filtering (e.g. by EPC product code or other arbitrary criteria) and flow control to manage bursts of data otherwise flooding the backend. All activity can be logged in an audit database and defined events may be escalated.

DEFEATING COUNTERFEITING AND TAMPERING

The second major attack type against RFID systems has to do with cloning of tags. To achieve this the system must (1) guarantee data integrity (prove tag data has not been manipulated), (2) uniquely associate the tag data with the tag itself (prove the data has not been copied from the original tag to another tag) and (3) uniquely associate the data with the real-world object it is describing (prove the tag has not been peeled off the original object and relabeled to another one). One possible solution to these problems is the use of cryptographic hash functions or digital signatures. The advantage of this approach is that no resources are required on the tag to calculate or verify these signatures. All computationally expensive operations can be carried out in a dedicated appliance inspecting all inbound and outbound RFID traffic.

To detect data manipulations a digital signature (e.g. based on a Hash Message Authentication Code or HMAC) must be calculated over the data stored on the tag. The signature should not only take the writable tag data into account but also a data field that is unique to the tag (e.g. an unchangeable tag ID as is common for ISO 15693 tags and also available for certain EPC tags). Ideally, the digital signature should also be associated with a third piece of data, some kind of "finger-print" or object ID, which is unique to the object or product the tag data is describing (e.g. spectrographic analysis, precise weight or form factor). By calculating a single digital signature over the aggregate of these three data sets (tag data, tag ID and object ID) we can establish data integrity and also establish a trusted link between tag data and tag as well as tag data and object, addressing all three problems described above at the same time.

If there is sufficient user memory available on the tag, the digital signature can be stored directly on the tag (Offline Crypto System, Solution 2) or if this is not the case the signature can be stored in a separate signature database (Online Crypto System, Solution 3).

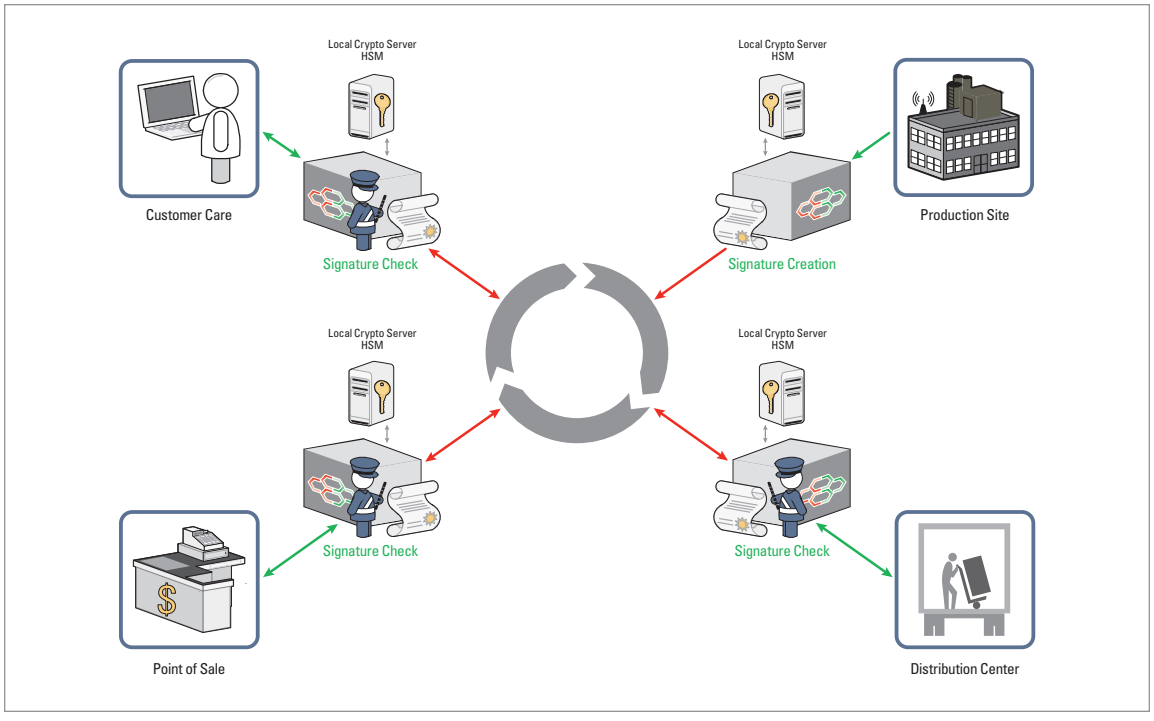

Diagram 3: Offline Crypto System

DEFEATING COUNTERFEITING AND TAMPERING (continued)

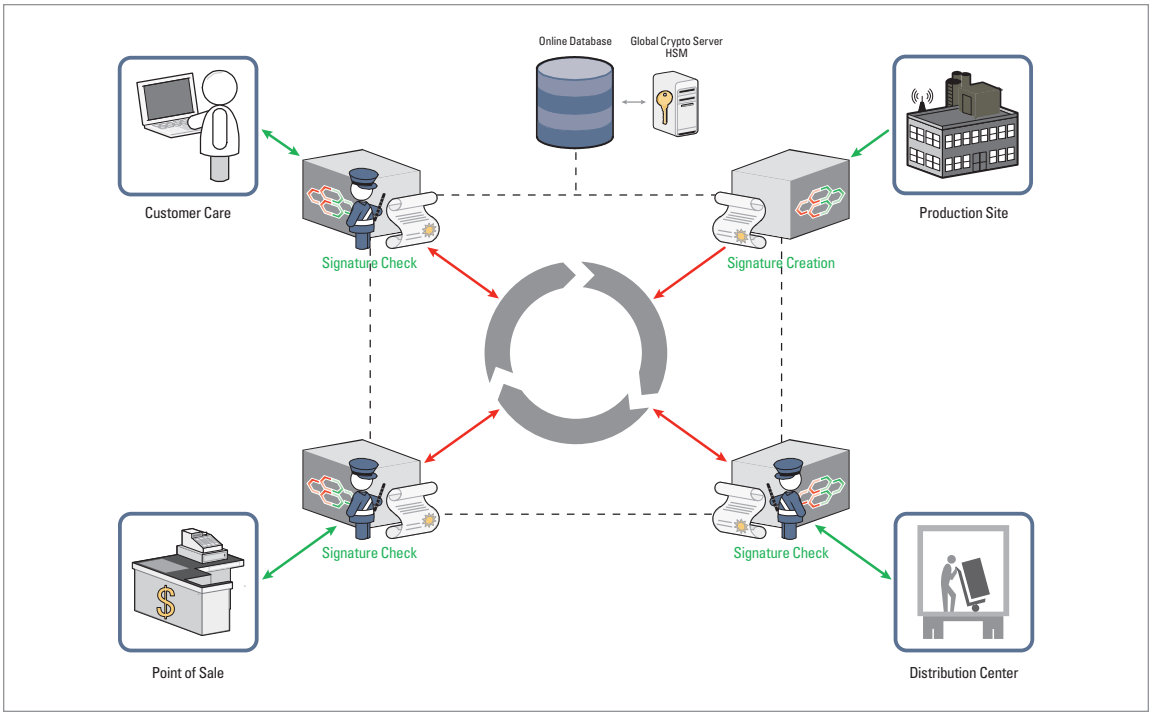

Diagram 4: Online Crypto System

The first time the tag is written, e.g. at the manufacturing site, the digital signature is calculated. This can be achieved using an architecture similar to the one shown in diagram 3: An appliance deployed right in front of the reader intercepts the write operation initiated by the back-end, calculates the signature and either injects the signature into the tag write process or stores it in a separate database.

Once the product has been shipped and arrives at a distribution center or warehouse, a second installation of the same appliance now reads the tag along with its associated signature. The solution follows the same algorithm to re-calculate the signature and finally compares the two signatures. If they are identical, tag data integrity has been established. If they are not, the tag data has been cloned or modified.

Both, the online and offline crypto solutions implement the same digital signature / HMAC algorithms which all rely on a shared secret key. This requires careful key manage management: The secret key could be stored either directly inside the RFID security appliance in a protected memory area (low-end solution), in a trusted platform module (TPM), providing a higher level of security, or using a hardware security module (HSM) as a dedicated crypto server providing the highest level of security.

The Offline Crypto System (Solution 2) requires that each instance of the RFID security appliance has its own key storage; the Online Crypto System (Solution 3) on the other hand performs all key management at a single location. This is generally considered to be more secure but it requires a reliable VPN communication link between each instance of the RFID security

DEFEATING COUNTERFEITING AND TAMPERING (continued)

appliance in the field and the global crypto server which is often impractical.

NeoCatena provides solutions suitable for implementing an Online or Offline Crypto System as outlined above offering various options for HMAC algorithms and key management. The solution is implemented as a dedicated appliance intercepting all data communication between reader and back-end initiating the necessary signature calculation, storage and validation operations.

Any security solution should be tailored to the specific requirements of a given system, so that the security level provided is adequate for the application at hand. An RFID security solution is no exception from this rule. For example, a basic protection level can be provided by implementing Solution 1 alone, while a higher level of security is offered by an implementation consisting of Solution 1+2 or Solution 1+3.

CONCLUSION

RFID enabled applications have significant benefits over applications based on barcodes, but the technology also introduces a number of new business risks. For example, a thief might modify or clone RFID tags at the point of sale to trick an automated checkout system (cyber shop lifting). Attackers might use specifically prepared tags to attempt injecting viruses or other Malware into backend systems to exploit existing software vulnerabilities and as a result disrupt business continuity. Electronic passports or other RFID based ID cards may be tampered with to gain illegal access.

With on-tag security being difficult to implement and too expensive for low-cost applications we have to look for off-tag security solutions to defeat the main business risks in RFID: Counterfeiting and attacks against the back-end connected to the edge of the network. NeoCatena meets these requirements with RF-Wall, a dedicated security appliance that can identify counterfeit tags as well as tags containing RFID Malware. RF-Wall is deployed directly behind the RFID Reader and in front of Middleware and Edge Servers to inspect tag data before it can reach the backend and thus shield the backend from these risks.

NeoCatena Networks Inc.

>> Next Generation RFID Security >>

440 N. Wolfe Rd, Sunnyvale, CA 94085
t 408.524.3116 >> f 650.472.8941
www.neocatena.com

ABOUT NEOCATENA

NeoCatena Networks supplies security solutions designed to minimize business risks inherent to IT technology and secure RFID applications across industries. The company was founded in 2006 by leading experts in the field of IT security with the goal to engineer products addressing key RFID security risks including attack, fraud, error and data leak. NeoCatena's technology is based on patent pending intellectual property that is key for any solution aimed at managing business risks inherent to RFID technology. NeoCatena is a privately owned company located in Sunnyvale, California