A study conducted by the Verizon Business RISK Team

2009 Data Breach Investigations Report

verizon business

2009 Data Breach Investigations Report

A study conducted by the Verizon Business RISK team.

For additional updates and commentary, please visit http://securityblog.verizonbusiness.com.

AUTHORS:

Wade H. Baker
Alex Hutton
C. David Hylender
Christopher Novak
Christopher Porter
Bryan Sartin
Peter Tippett, M.D., Ph.D.
J. Andrew Valentine

CONTRIBUTORS:

Thijs Bosschert
Eric Brohm
Calvin Chang
Ron Dormido
K. Eric Gentry
Mark Goudie
Ricky Ho
Stan S. Kang
Wayne Lee
Jelle Niemantsverdriet
David Ostertag
Michael Rosen
Enrico Telemaque
Matthijs Van Der Wel
Ben Van Erck
Members of the RISK Team
ICSA Labs

SPECIAL THANKS TO:

Janet Brumfield
Carl Grygiel
Hunter Montgomery

TABLE OF CONTENTS

2009 Data Breach Investigations Report

A study conducted by the Verizon Business RISK team

Executive Summary

2008 will likely be remembered as a tumultuous year for corporations and consumers alike. Fear, uncertainty, and doubt seized global financial markets; corporate giants toppled with alarming regularity; and many who previously lived in abundance found providing for just the essentials to be difficult. Among the headlines of economic woes came reports of some of the largest data breaches in history. These events served as a reminder that, in addition to our markets, the safety and security of our information could not be assumed either.

The 2009 Data Breach Investigations Report (DBIR) covers this chaotic period in history from the viewpoint of our forensic investigators. The 90 confirmed breaches within our 2008 caseload encompass an astounding 285 million compromised records. These records have a compelling story to tell, and the pages of this report are dedicated to relaying it. As with last year, our goal is that the data and analysis presented in this report prove helpful to the planning and security efforts of our readers. Below are a few highlights from the report:

Who is behind data breaches?

74% resulted from external sources (+1%).

20% were caused by insiders (+2%).

32% implicated business partners (-7%).

39% involved multiple parties (+9%).

Closely resembling the stats from our 2008 report, most data breaches continue to originate from external sources. Though still a third of our sample, breaches linked to business partners fell for the first time in years. The median size of breaches caused by insiders is still the highest but the predominance of total records lost was attributed to outsiders. 91 percent of all compromised records were linked to organized criminal groups.

How do breaches occur?

67% were aided by significant errors (<>).

64% resulted from hacking (+5%).

38% utilized malware (+7%).

22% involved privilege misuse (+7%).

9% occurred via physical attacks (+7%).

In the more successful breaches, the attacker exploited some mistake committed by the victim, hacked into the network, and installed malware on a system to collect data. 98 percent of all records breached included at least one of these attributes. Unauthorized access via default credentials (usually third-party remote access) and SQL injection (against web applications) were the top types of hacking. The percentage of customized malware used in these attacks more than doubled in 2008. Privilege misuse was fairly common, but not many breaches from physical attacks were observed in 2008.

What commonalities exist?

69% were discovered by a third party (-6%).

81% of victims were not Payment Card Industry (PCI) compliant.

83% of attacks were not highly difficult (<>).

87% were considered avoidable through simple or intermediate controls (<>).

99.9% of records were compromised from servers and applications.

Only 17 percent of attacks were designated to be highly difficult, yet they accounted for 95 percent of the total records breached. So, while hackers prefer soft targets, they do seem to know where best to apply the pressure when motivated. Most of these incidents do not require difficult or expensive preventive controls; mistakes and oversight hinder security efforts more than a lack of resources. 81 percent of organizations subject to PCI DSS had not been found compliant prior to the breach. Nearly all records in 2008 were compromised from online assets. As with last year's report, the majority of breaches are discovered by a third party.

Where should mitigation efforts be focused?

Some will recognize three of these five recommendations as carryovers from our previous report. This is intentional. We simply could not convince ourselves to remove them just to avoid reiteration. In fact, a fresh look and further consideration is warranted.

The best defense against data breaches is, in theory, quite simple—don't retain data. Since that is not realistic for many organizations, the next best thing is to retain only what is required for business or legal reasons, to know where it lives and flows, and to protect it diligently.

The majority of breaches still occur because basic controls were not in place or because those that were present were not consistently implemented across the organization. If obvious weaknesses are left exposed, chances are the attacker will exploit them. It is much less likely that they will expend the time and effort if none are readily apparent.

As a specific extension of this, we felt it necessary to call out several tried and true controls based on our 2008 case data. A very large proportion of attackers gain access to enterprise networks via default, shared, or stolen credentials. Furthermore, organizations seem to have little visibility into this problem. It's certainly best to prevent such incidents in the first place, but a second line of defense is to review accounts for signs of abuse or anomalies. SQL injection was also an oft-used means of breaching corporate data last year. Secure development, code review, application testing, etc. are all considered beneficial in light of this finding.

Whatever the sophistication and aggressiveness of attacks, the ability to detect a breach when it occurs is a huge stumbling block for most organizations. Whether the deficiency lies in technology or process, the result is the same—during the last five years, few victims discover their own breaches. Fewer still discover them in a timely manner.

- ✔ Ensure essential controls are met.
- ✔ Find, track, and assess data.
- ✔ Collect and monitor event logs.
- ✔ Audit user accounts and credentials.
- ✔ Test and review web applications.

Methodology

The underlying methodology used in this report remains unchanged from the previous year. All results are based on firsthand evidence collected during data breach investigations conducted by Verizon Business from 2004 to 2008. The 2008 caseload is the primary analytical focus of the report, but the entire range of data is referenced extensively throughout. Though the Investigative Response (IR) team works a variety of engagements, only those involving a confirmed breach are included in this data set. To help ensure reliable and consistent input, all investigators use the same standardized tool to record case data and other relevant details. This information is then submitted to other members of the RISK team for further validation and analysis.

Beyond this, there are a few notable differences and additions with respect to the 2009 Data Breach Investigations Report. Whereas the 2008 report reached back across four years of cases in one massive data collection effort, this data set was assembled periodically throughout the year. Investigators were able to enter information at the close of a case while it was still fresh in their minds. This shift from historic to ongoing collection allows for more detail on existing data points and opens the door to new areas of study. We hope these additions enhance the value and utility of this report to the research and practitioner communities.

Most of the statistics presented in this report refer to the percentage of cases, the percentage of records breached, or simply the number of cases. The "percentage of records" statistic is new this year and gives a sometimes different but always insightful view of the data. Because of the potentially misleading nature of assigning percentages to small samples, the raw number of cases is used anytime we discuss a subsample within the caseload. For instance, evidence of malware was found in 38 percent of cases, and in the several pages dedicated to these attacks, all figures show integers. Captions and legends should aid proper interpretation.

Whereas the 2008 report reached back across four years of cases in one massive data collection effort, this data set was assembled periodically throughout the year. This shift from historic to ongoing collection allows for more detail on existing data points and opens the door to new areas of study.

We would like to reiterate that we make no claim that the findings of this report are representative of all data breaches in all organizations at all times. These statistics are based solely upon our caseload and any conclusions or inferences we make are drawn from this sample. Although we believe many of these results to be appropriate for generalization, bias undoubtedly exists. Even so, there is a wealth of information here and no shortage of valid and clear takeaways. As with any study, readers will ultimately decide which findings are applicable within their organization.

Finally, it is important to note that Verizon Business is committed to maintaining the privacy and anonymity of Investigative Response clients. Once the investigator records and submits case metrics, this information is sanitized and the client's name is removed from the records. The central repository of case data contains no information that would enable one to ascertain a client's identity. Furthermore, the statistics within this report are always presented in aggregate; individual records are never the focus of analysis.

State of Cybercrime, 2009

Before delving into the statistics and analysis presented in our 2009 report, we thought it a good idea to update the "Primer on Cybercrime" originally presented in the 2008 DBIR. This brief section attempts to put some context around the data and highlight important aspects of the continuing evolution of cybercrime around the world. One may doubt that the cybercrime market could change much over a single year, but one need only consider global financial markets in 2008 to realize that any market system can change and, at times, change swiftly. As the cybercrime market evolves, attackers, targets, and techniques do as well.

The potential value of engaging in cybercrime would not exist without a market for stolen data. As with any legitimate market system, the unit value of goods and services fluctuates with supply and demand. Massive exposures of magnetic-stripe data in recent years (hundreds of millions in our caseload alone) have effectively flooded the information black market, saturating it with "dumps," or credit card magnetic stripe sequences sufficient for counterfeit. This market saturation has driven the price down to a point where magnetic-stripe information is close to worthless. The value associated with selling stolen credit card data have dropped from between $10 and $16 per record in mid-2007 to less than $0.50 per record today.*

As supply has increased and prices fallen, criminals have had to overhaul their processes and differentiate their products in order to maintain profitability. In 2008, this was accomplished by targeting points of data concentration or aggregation and acquiring more valuable sets of consumer information.

As supply has increased and prices fallen, criminals have had to overhaul their processes and differentiate their products in order to maintain profitability. In 2008, this was accomplished by targeting points of data concentration or aggregation and acquiring more valuable sets of consumer information. The big money is now in stealing personal identification number (PIN) information together with associated credit and debit accounts. Thus, we saw an explosion of attacks targeting PIN data in the previous year. These PIN-based attacks hit the consumer much harder than typical signature-based counterfeit attacks. This is because PIN fraud typically leads to cash being withdrawn directly from the consumer's account—whether it be a checking, savings, or brokerage account. Furthermore, PIN fraud typically places a larger share of the burden upon the consumer to prove that transactions are fraudulent. This makes the recovery of lost assets more difficult than with standard credit-fraud charges.

The higher value commanded by PIN data has spawned a cycle of innovation in attack methodologies. Criminals have reengineered their processes and developed new tools—such as memory-scraping malware—to steal this valuable commodity. This has led to the successful execution of complex attack strategies previously thought to be only theoretically possible. As a result, our 2008 caseload is reflective of these trends and includes more targeted, cutting edge, complex, and clever cybercrime attacks than seen in previous years..

*Figures based on data collected as part of Verizon Business underground intelligence operations.

Results and Analysis

The Verizon Business IR team worked well over 150 forensic engagements in 2008. Of those, 90 were data compromise investigations in which a breach was confirmed. A number of these investigations were quite extensive and lengthy; a fact which contributed to the lower-than-average number of cases worked this year. Though fewer, these 90 held their own; the total number of records breached across our 2008 caseload—more than 285 million—exceeded the combined total from 2004 to 2007.

At the time of this writing, about a third of the breaches investigated by our team last year are publicly disclosed. More, especially those toward the end of the year, are likely to follow. Others will likely remain unknown to the world as they do not fall under any legal disclosure requirements.

Roughly 20 percent of 2008 cases involved more than one breach. That is to say, multiple distinct entities or locations were individually compromised as part of a single case. Amazingly, nearly half of our caseload was comprised of different sets of interrelated incidents. Quite often the same individual(s) committed the attack. Other times, there was a shared connection (literally) between the victims and a common third party that experienced a breach. Still others were linked through some kind of common application, identical attack patterns, and the like.

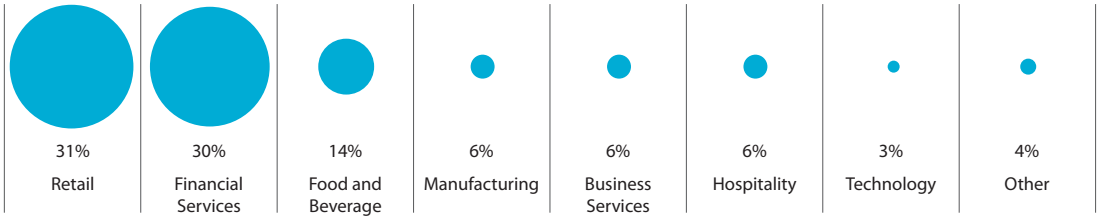
These 90 cases along with those worked between 2004 and 2007 form the basis of all results and analysis within this report.

Demographics

As with last year's report, data breaches affected a wide array of organizations in 2008. These are categorized according to the industry groups presented in Figure 1. Claiming nearly a third of all breaches, retail continues to be the most frequently affected industry. Food and beverage establishments, second-most common in the 2004 to 2007 data set, dropped in both proportion (20 percent to 14 percent) and position (now third place) in 2008. The major gainer in 2008 was financial services, which doubled in terms of caseload percentage to 30 percent.

The increase of data breaches in the financial sector is indicative of recent trends in cybercriminal activity highlighted in the "State of Cybercrime" section. As will be discussed throughout this report, financial services firms were singled out and fell victim to some very determined, very sophisticated, and—unfortunately—very successful attacks in 2008. This industry accounted for 93 percent of the over 285 million records compromised. This finding reflects a few very large breaches

Figure 1. Industries represented by percent of breaches

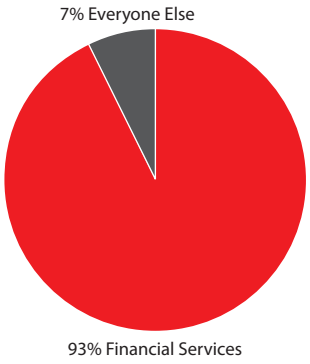


Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Demographics" post.

investigated by our IR team in the past year. Though few in number, they dominate all percentage of records statistics discussed throughout this report.

Figure 2. Industries represented by percent of records



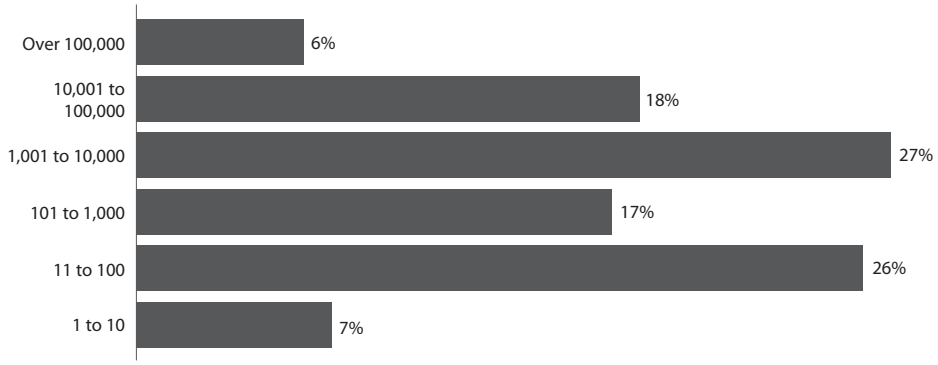
Beyond these top three industry groups, a smattering of others filled out the remaining quarter of cases. Manufacturing and business services (which includes a few media, marketing, consulting, and legal firms) and hospitality each accounted for 6 percent of the caseload. Technology firms, which made up 13 percent of our 2004 to 2007 cases, were comparatively less represented in 2008. We view this difference to be more reflective of our sample than a broader trend.

The number of investigations handled by our IR team outside the United States rose to over one-third of our caseload in 2008. In addition to extensive investigations across the United States, many breaches hit organizations in Canada and Europe while casework demands continued to grow in Brazil, Indonesia, the Philippines, Japan, and Australia. As attackers continue to pursue soft targets internationally, concern in emerging economies will rise as well, especially with respect to consumer data.

The distribution of organizational size looks very similar to the previous data set. Per Figure 3, data thieves seem to show no partiality between larger enterprises and smaller establishments. Though not always the case, criminals typically initiate attacks based on perceived value of the data and convenience rather than victim characteristics such as size.

One final point of interest deserves mention before concluding this section. A newly added line of inquiry for 2008 found that 13 percent of organizations in our caseload had recently been merged or acquired. It's difficult to draw a conclusion from this

Figure 3. Number of employees by percent of breaches



statistic or assign any significance to it—yet the potential effect of such changes on the likelihood of suffering a breach is worth considering.

Mergers and acquisitions bring together not only the people and products of once separate organizations but their technology environments as well. Integration rarely happens overnight or without a hitch. Technology standards are sometimes set aside for the sake of business expediency. This introduction of variance into the IT operating environment may serve to increase the risk of compromise. Furthermore, businesses preparing for sale may find reducing operating expenses—including cutbacks to IT and security spending—a convenient way to help the balance sheet at the time of sale. Finally, new ownership may alter (by mandate or by culture) the acquired organization's tolerance for information risk.

All this, of course, is speculation and cannot be proven or disproven (or even tested) without additional information. We added it to our case metrics with the idea that it might reveal something more substantial over time and we will continue to record and report it.

Sources of Data Breaches

Similar to cases conducted in the physical realm, one of the primary objectives during a computer forensics investigation is to identify those responsible for the crime. Because perpetrators often return to the scene, knowing the source of a breach can be essential to its containment. At a high-level, security incidents originate from one or a combination of the following sources:

External: External threats originate from sources outside the organization. Examples include hackers, organized crime groups, and government entities, as well as environmental events such as weather and earthquakes. Typically, no trust or privilege is implied for external entities.

Internal: Internal threat sources are those originating from within the organization. This encompasses human assets—company executives, employees and interns—as well as other assets such as physical facilities and information systems. Most insiders are trusted to a certain degree and some, IT administrators in particular, have high levels of access and privilege.

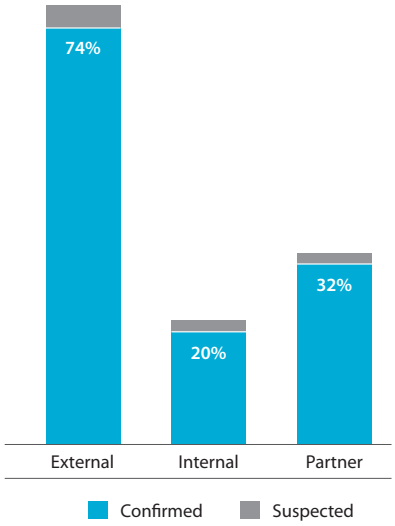
Partner: Partners include any third party sharing a business relationship with the organization. This value chain of partners, vendors, suppliers, contractors, and customers is known as the extended enterprise. Information exchange is the lifeblood of the extended enterprise, and, for this reason, some level of trust and privilege is usually implied between business partners.

Results from 600 incidents over five years make a strong case against the long-abiding and deeply held belief that insiders are behind most breaches.

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Sources of Data Breaches" post.

Figure 4. Sources of breaches by percent of breaches



If evidence reveals that any of these played a significant and active role in the breach, it is marked as a source. While there is some room for interpretation in "significant and active," investigators do follow a set of consistent guidelines. For instance, an insider that deliberately steals proprietary information from their employer is clearly an "internal" breach. We also consider insiders partially responsible when their actions, though unintentional, either directly cause or contribute to the breach. Picking up malware while browsing that is later used by an external attacker to gain unauthorized access is an example of this. We *do not* consider it an internal source when an insider's inaction (i.e., oversight, failure to follow-through on procedures, decision to not implement certain security measures, etc) allows or aids a breach. The distribution of breach sources in 2008 is presented in Figure 4. The results are quite similar to that of the 2004 to 2007 data set and continue to challenge some of the prevailing wisdom in the security community with regard to the origins of data breaches.

Prior to further discussion of these results, it's worth clarifying two points of potential confusion. First, it is no mistake that the values in Figure 4 sum to more than 100 percent, as many breaches involve multiple parties. Figure 5 below illustrates the distribution of breach sources to highlight this fact. Second, we want to be clear that these findings relate specifically to the occurance (or likelihood) of security breaches leading to data compromise within our caseload—not attacks, not impact, not general security incidents, and not risk. We observed some rather strong reactions to this finding after last year's report, and it was apparent that at least some of the discussion had more to do with terminology than the actual results.

The majority of data breaches continue to originate from sources outside the victim. In 2008 Verizon Business encountered nearly the same percentage (74 percent) of confirmed external breaches as our combined 2004 to 2007 caseload. Furthermore, this statistic remains remarkably consistent over the five-year period of this study. Based on these results, it seems unwise to downplay the threat posed by outsiders.

Figure 5. Single vs. multiple breach sources by percent of breaches

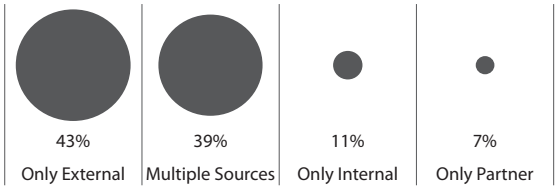
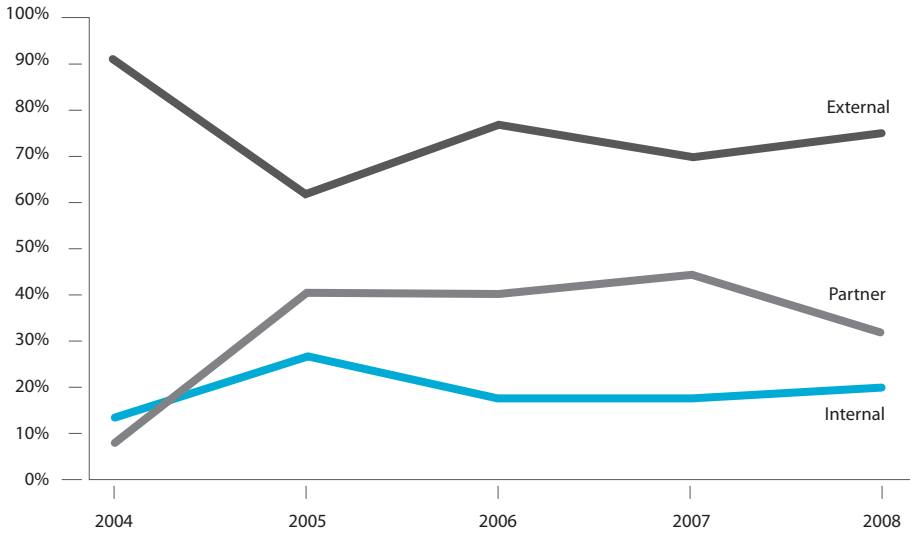


Figure 6. Breach sources over time by percent of breaches



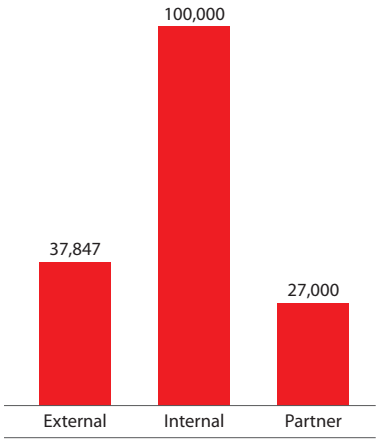
Insiders, on the other hand, are behind the lowest proportion (20 percent) of breaches in our caseload for four years running. Figure 5 shows that only about half of these (11 percent of all breaches) were committed by an insider acting alone. The remainder of the breaches tied to insiders mostly involved employees as unwitting participants in the crime through errors and policy violations. It is true that these results are based upon our caseload—which is consumer data-heavy—and may not be reflective of all data breaches. Perhaps insiders are more apt to target other types of data such as intellectual property. It is also true that many insider crimes may never be detected, though one would think any breach causing material harm would eventually be noticed. It is also feasible they are more likely handled internally. At any rate, results from 600 incidents over five years make a strong case against the long-abiding and deeply held belief that insiders are behind most breaches.

The number of breaches linked to business partners continues to land between external and internal sources but did drop 12 percent in 2008. Interpreting this decline is difficult as it is doubtful that huge strides were made in the effort to reduce partner-facing risk. It is more likely related to the lower proportion of food and beverage and retail cases within our 2008 caseload. Readers of last year's supplemental report may remember that those two industries exhibited high percentages of partner-related breaches (particularly food and beverage at 70 percent or more). In contrast, this year's results show that criminals appear to be directly targeting victims that offer a bigger payout. The "end around" maneuver via trusted partner infrastructure does not seem to be the vector of choice in these attacks. Nevertheless, breaches involving partners are still quite common and account for over one-third of cases if both confirmed and suspected cases are counted. Any other difference from past data that cannot be explained due to caseload composition is likely insignificant statistical variation.

Breach Size by Source

Figure 7 shows the median* number of records compromised per event for each threat source. As a reminder, we do not assert that the full consequences of a breach are limited to the number of records exposed; we use this statistic merely as a measurable indicator of the overall impact.

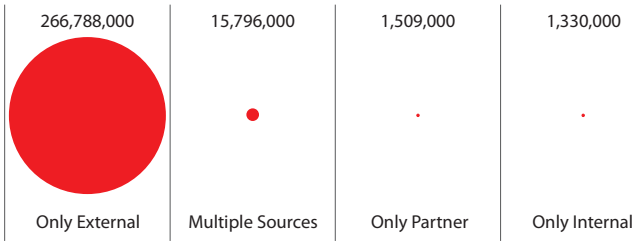
Figure 7. Median number of records compromised per breach



Insider breaches (individually) continue to be much more damaging than those caused by other sources though the difference between them is not to the extent observed across our 2004 to 2007 caseload. One of the more interesting changes is that outsiders compromised more records per incident than partners. This shift is attributable to several very large breaches investigated in 2008 which were perpetrated by outsiders. A comparison of the median value provided at right (37,847) with the mean** (5,651,067) gives an appreciation for the dramatic skew that exists within the data set with respect to the size of external breaches. This is one of several reasons why we use the median as the preferred measure of central tendency when analyzing these incidents. Figure 8 provides a striking view of the size and dominant nature of external breaches last year.

At this point, those familiar with our pseudo risk calculation (likelihood x impact) and its result in the last report may suspect that it will yield a different outcome this year. That instinct would be correct. Case results from 2008 find that outsiders represent the greatest risk for data compromise, followed closely by insiders and then partners. This presents a pattern exactly opposite from what was depicted in our 2004 to 2007 data set. Does this mean that the fundamental nature of information risk experienced a profound metamorphosis last year? It is doubtful; keep in mind that risk is probabilistic and best understood over time with multiple measurements. Though few in number, several large breaches were enough to tip the scales in the direction of outsiders as the dominant source in 2008.

Figure 8. Total records compromised by source



*The middle value in an ascending set of numbers

**The average of a set of numbers

Table 1. Pseudo risk calculation

Source	Likelihood	Impact (number of records)	Risk (pseudo)
External	**74%**	37,847	28,175
Internal	20%	**100,000**	20,000
Partner	32%	27,000	8,700

External Breach Sources

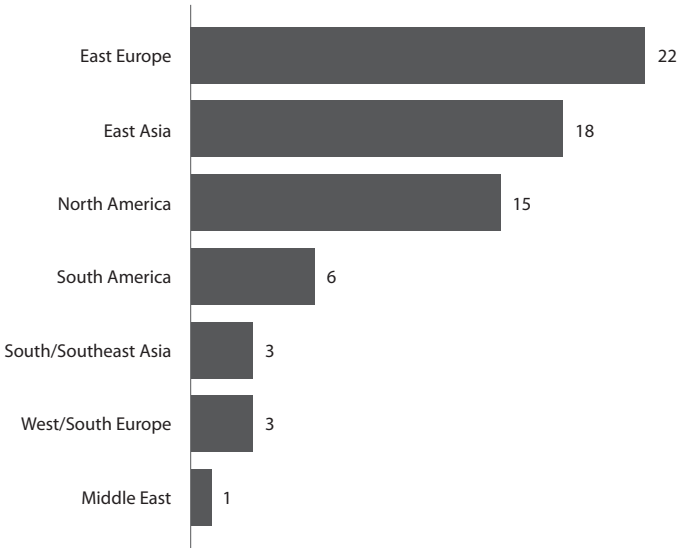
The true geographic origin of an attack is difficult to pinpoint with certainty. This determination is predicated upon the source IP address, which is often unreliable for many reasons. Even so, additional validation is gained through common elements between cases, correlative fraud patterns, information provided by other Verizon Business departments, and collaboration with law enforcement agencies. The geographic distribution of external data breach sources is shown in Figure 9.

Though in slightly different order, Eastern Europe, East Asia, and North America remain at the top of the list in 2008. In fact, these regions are even more dominant, accounting for 82 percent of all external attacks. By comparison, 59 percent of breaches between 2004 and 2007 originated from these regions. Eastern Asia (up 15 percent) and Eastern Europe (up 9 percent) are most responsible for the change.

Though it's tempting to pander to hype surrounding state-sponsored attacks from Asia, we find no evidence to support the position that governments are a significant agent of cybercrime. We do have a great deal of evidence that malicious activity from Eastern Europe is the work of organized crime. This is further seen from Figure 10 which categorizes external entities into familiar types rather than by region.

That nearly two-thirds are "not traced to a specific entity other than IP" is the result of several factors. Sometimes we are unable to do so. Other times the victim decides it is not worth the additional time and expense. In most cases, the immediate need with respect to the IP address is in containing the breach rather than rooting out the

Figure 9. Location of attacking IP(s) by number of breaches



entities responsible. In those instances when an attempt is made to trace the IP to a specific entity, we work with law enforcement personnel. As seen in the chart, the trail often leads to members of known organized crime outfits. What is not evident from Figure 10 is the astounding statistic that 91 percent of all compromised records in 2008 was attributed to organized criminal activity. On the brighter side, we are happy to report that these efforts with law enforcement led to arrests in at least 15 cases (and counting) in 2008.

Figure 10. Catagories of external breach sources by number of breaches

Internal Breach Sources

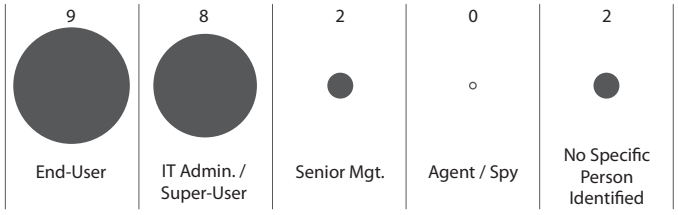
Several broad classifications of insiders are presented in Figure 11 along with the percentage of incidents attributed to each. 2008 results are similar to the 2004 to 2007 data set. End-users and IT administrators continue to be the culprits behind most breaches. This finding for IT administrators is not surprising; higher privileges afford greater opportunity and temptation for abuse. At the same time, the results for incidents perpetrated by end-users serve to remind us that internal breaches are not solely dependent on privileges or administrative credentials. Though our metrics do include options for part-time and temporary workers, our caseload included none.

Of all insider cases in 2008, investigators determined about two-thirds were the result of deliberate action and the rest were unintentional. While it's tempting to infer that administrators acted more deliberately and maliciously than end-users and other employees, the evidence does not support this conclusion. The ratio was roughly equal between them. It is worth noting that both cases involving senior management were the result of deliberate action which was taken after the person was terminated. We also noticed several other breaches in the caseload were perpetrated by recently terminated employees. The majority was administrators, but a few cases involved end-users as well. With respect to breaches caused by recently terminated employees, the following two scenarios were observed:

- Employee was terminated and his/her account was not disabled in a timely manner.
- Employee was notified of termination but was allowed to "finish the day" unmonitored and with normal access/privileges.

This obviously speaks to the need for termination plans that are timely and encompass all areas of access (decommissioning accounts, disabling privileges, escorting terminated employees, etc.).

Figure 11. Catagories of internal breach sources by number of breaches



Partner Breach Sources

The majority of breaches involving a business partner was the result of third-party information assets and connections being compromised and used to attack the victim's systems. This statistic increased substantially in 2008. (For the sake of reference, it was slightly over half of 2004 to 2007 cases.) In the large majority of cases, it was the lax security practices of the third party that allowed the attack. It should not come as a surprise that organizations frequently lack measures to provide visibility and accountability for partner-facing systems.

Figure 12. Categories of partner breach sources by number of breaches

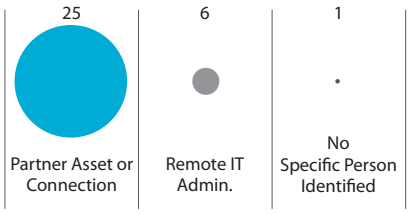


Figure 12 also reminds us that not all data breaches within the extended enterprise are unintentional. Rising to a slightly higher proportion last year, six instances of deliberate malicious action by third-party remote administrators were observed. One of these individuals had been recently terminated.

After last year's report, we had many inquiries regarding the nature of the relationship between the victim (client) and the partner. We attempted to capture this information during 2008 investigations. This new information found that most of these breaches dealt with a partner who administered victim-side assets. For retail and food and beverage organizations, this was almost always a vendor supporting a point-of-sale (POS) system. We also noted several instances where the partner had user-level access to the victim's systems or regularly exchanged data with the victim. Only one case involved a partner physically handling or transporting victim assets. Interestingly, our caseload included zero instances where the partner hosted the victim systems.

In the large majority of cases, it was the lax security practices of the third party that allowed the attack. It should not come as a surprise that organizations frequently lack measures to provide visibility and accountability for partner-facing systems.

Threat and Attack Categories

Anyone responsible for safeguarding corporate information assets knows there are countless ways in which sensitive information will find its way into the wrong hands. Though sometimes one-dimensional, data breaches are more often the result of a series of intertwined and orchestrated events. Examining the frequencies and trends surrounding these scenarios is essential to protection efforts and is the purpose of this section.

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Threat and Attack Categories" post.

Though very specific attack details are noted during an investigation, all possibilities fit somewhere within the seven top-level threat categories listed in the figure below. Figure 13 records the prevalence of each as causing or contributing to data breaches investigated by Verizon Business in 2008 (black bars). Since most incidents involve events spanning several categories, the percentages sum to well over 100 percent. Also depicted is the percentage of total compromised records ascribed to each category (red bars).

The results for 2008 cases look very similar to those of the 2004 to 2007 data set (see Figure 14 below for a time series chart of these results). The Deceit and Physical categories switched places, but all others remained in order. Furthermore, Hacking and Malware continue to dominate the caseload. Error is seldom the proximate cause of a breach, but it is very often a factor contributing to or enabling a successful attack. From Figure 13, one can deduce the stereotypical breach scenario: the attacker takes advantage of some mistake committed by the victim, hacks into the network, and installs malware on a system to collect data. As evidenced by the red bars in Figure 13, this is especially true for large breaches. The following sections provide a more in-depth examination of each threat category.

Figure 13. Threat categories by percent of breaches (black) and records (red)

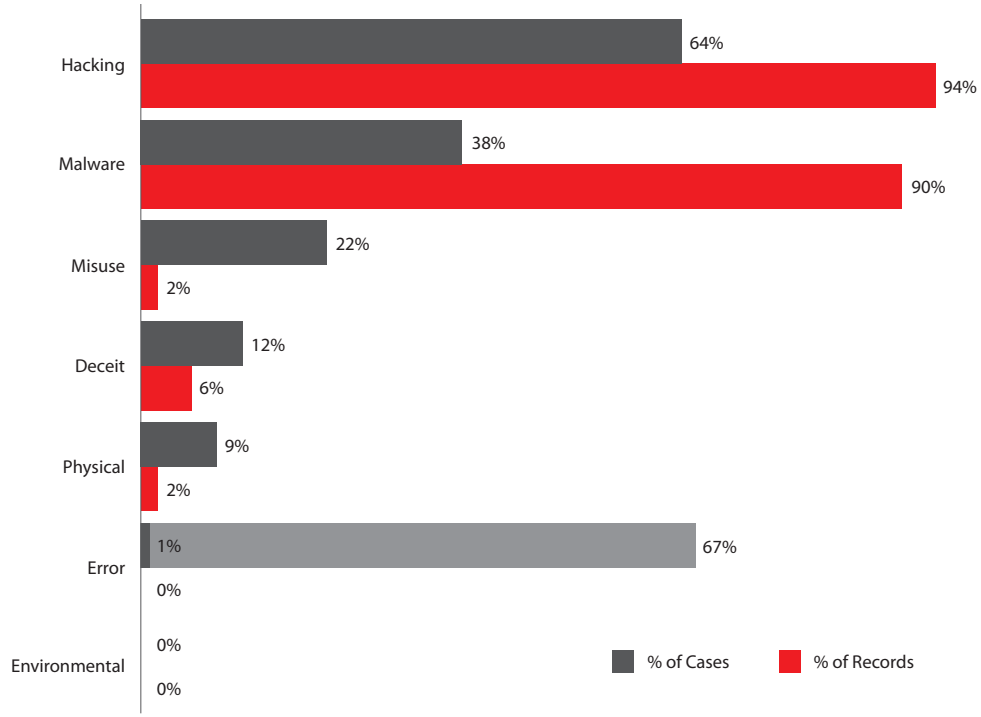
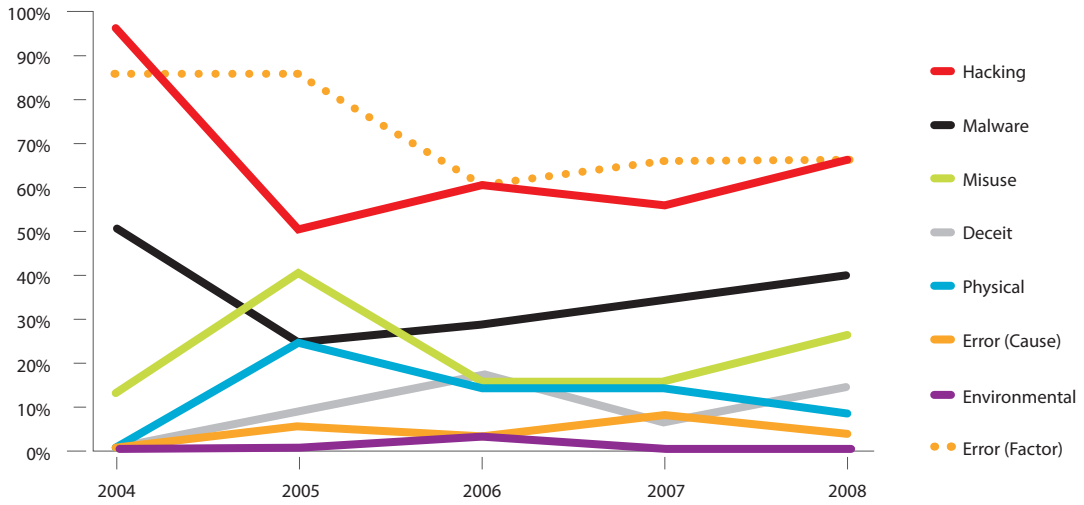


Figure 14. Threat categories over time by percent of breaches



Hacking and Intrusion

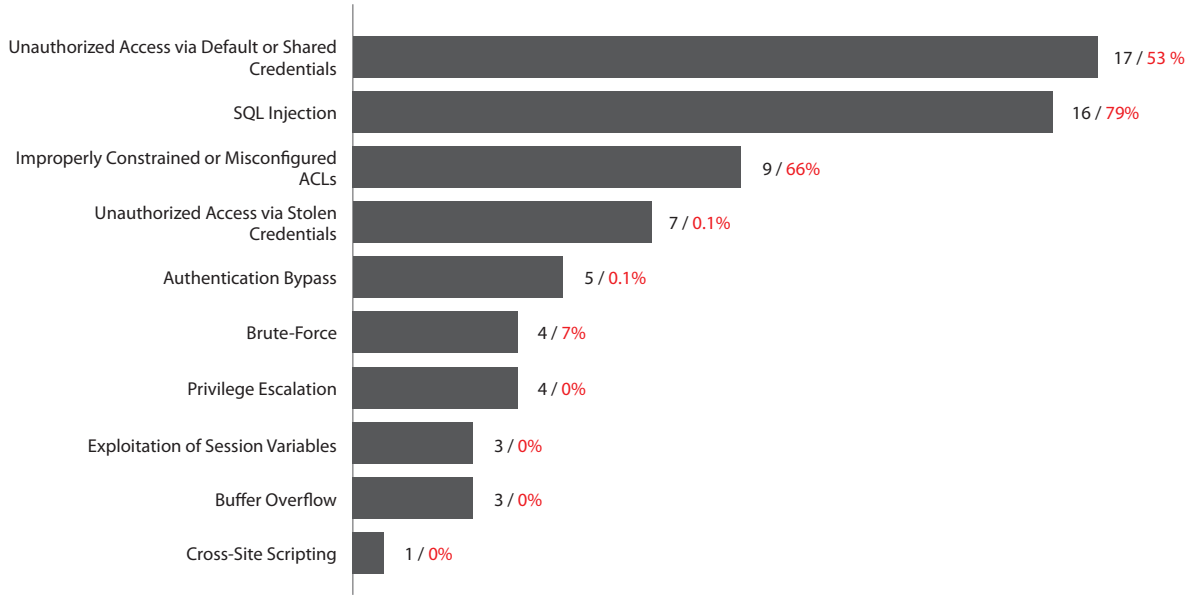
In terms of malicious action against information systems, hacking is the leading cause of data breaches for the fifth year running. Since hacking is less subject to the constraints that limit other attack methods (i.e., physical proximity, human interactions, special privileges), this is not unexpected. Additionally, many tools are available to help automate and accelerate the attack process, which keeps the cost of attack relatively low for the criminal. Those familiar with attack classification methodologies will know that the library of hacking and intrusion techniques is quite extensive. For 2008, we expanded our IR case metrics to provide more detail around this prevalent and potent threat category. Figure 15 reveals the types of hacking observed by Verizon Business during breach investigations in the last year.

From the chart, it is evident that many intrusions exploit the basic (mis)management of identity. Unauthorized access via default, shared, or stolen credentials constituted more than a third of the entire Hacking category and over half of all compromised records. It is particularly disconcerting that so many large breaches stem from the use of default and/or shared credentials, given the relative ease with which these attacks could be prevented. Readers may wonder why default and shared credentials are lumped together, as these categories seem to represent two different problems. The answer is that these issues were frequently found in tandem. We investigated an entire series of cases in which multiple organizations within the same industry all suffered breaches within a very short timeframe. It didn't take long to figure out that each used the same third-party vendor to remotely manage their systems. Unfortunately, that vendor neglected to change the default username and password—*and* used the same credentials across multiple clients.

Similarly disturbing are those breaches (and the high percentage of compromised records) traced to poor access control lists (ACLs). In more than a few cases, ACLs proved to be somewhat of a misnomer, leaving a wide-open door for the assailant to walk through unchallenged. Criminals will usually take the path of least resistance, and unfettered access fits that description quite well.

When hackers are required to work to gain access, SQL injection appears to be the uncontested technique of choice. In 2008, this type of attack ranked second in prevalence (utilized in 16 breaches) and first in the amount of records compromised (79 percent of the aggregate 285 million). At its most basic level, SQL injection attacks exploit a failure to properly validate user input. This seems especially common with custom-developed applications and web front-ends. In the absence of third-party OS and platform-specific vulnerabilities, criminals are aware of and exploiting weaknesses in application development processes. SQL injection has been a part of the security industry consciousness for years now, and some may wonder at its continued prevalence. Fixing vulnerable applications, however, can be challenging, costly, and time consuming, all of which contribute to a rather large and persistent attack surface. On top of this, SQL injection attacks are growing notably more sophisticated, especially for data compromise scenarios. It is often used to gain deeper access into systems and plant malicious software. Also noteworthy relative to hacking techniques is the infrequency with which more commonly known hacking techniques, such as buffer overflows, exploitation of session variables, and privilege escalation, appear in our data set.

Figure 15. Types of hacking by number of breaches (black) and percent of records (red)



Vulnerability Exploits

2008 continued a downward trend in attacks that exploit patchable vulnerabilities versus those that exploit configuration weaknesses or functionality. Only six confirmed breaches resulted from an attack exploiting a patchable vulnerability. The word "patchable" here is chosen carefully since we find that vulnerability does not have the same meaning for everyone within the security community. While programming errors and misconfigurations are vulnerabilities in the broader sense, lousy code can't always be fixed through patching and the careless administration patch has yet to be released. Furthermore, many custom-developed or proprietary applications simply do not have routine patch creation or deployment schedules.

For the six exploited vulnerabilities that had existing patches available, Table 2 shows how long the patch had been public at the time of the breach. The story is similar to that of the previous report; the interim between a patch's release and active exploits leading to data compromise is usually on the order of years. Vulnerabilities are certainly a problem contributing to data breaches, but patching faster is not the solution. This year's findings continue to support the idea that a patch deployment strategy focusing on coverage and consistency is far more effective at preventing data breaches than "fire drills" attempting to patch particular systems as quickly as possible.

Table 2. Patch availability at time of breach

Less than 1 month	0
1 to 3 months	0
3 to 6 months	0
6 to 12 months	1
More than 1 year	5

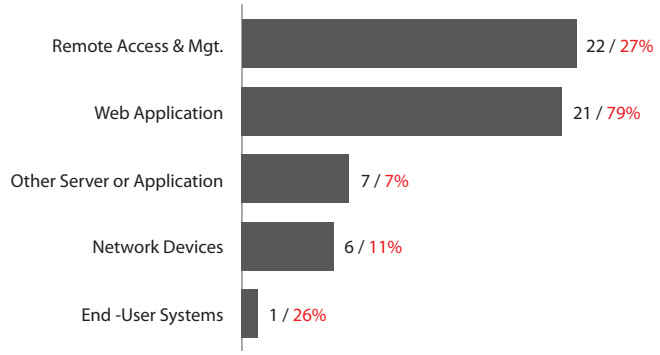
Attack Vector

Looking deeper into hacking activity, it is apparent that the bulk of attacks continues to target applications and services rather than the operating systems or platforms on which they run. Of these, remote access services and web applications were the vector through which the attacker gained access to corporate systems in the vast majority of cases. While network devices do sometimes serve as the avenue of attack, it was considerably less often in 2008.

In approximately four of 10 hacking-related breaches, an attacker gained unauthorized access to the victim via one of the many types of remote access and management software. Rather than for internal usage, most of these connections were provisioned to third parties in order to remotely administer systems. As discussed extensively in this and previous reports, the ultimate attacker is not typically the third party (although that certainly happens). More often, an external entity compromises the partner and then uses trusted connections to access the victim. From the victim's perspective the attacker appears to be an authorized third party, making this scenario particularly problematic. This is especially so when trusted access is coupled with default credentials.

Although web application attacks are one fewer in number than those against remote access services, they are responsible for a much larger number (79 percent) of breached records. Based on earlier discussion, it is not difficult to surmise that SQL injection is the predominate type of attack against this vector. Interestingly, the reason SQL injection is so successful is related to the scenario described above pertaining to remote access and management software. To function properly, a trust relationship must exist between web applications and back-end databases. In this sense, a request from the application is

Figure 16. Attack pathways by number of breaches (black) and percent of records (red)



similar to a request from a privileged administrator. The database obediently yields information requested by the application and cares not whether the command is valid or the result of an external attacker passing illegitimate strings. This is one of the primary reasons why encrypting databases is of limited effectiveness in preventing attack scenarios that do not involve physical theft or control of the system.

A much smaller percentage of hacks targeted routers, switches, and other network devices. This includes wireless networks, which continue to be a rare attack vector for data breaches. 2008 saw only a single instance in which a wireless network was exploited across our entire caseload (0.01 percent of all breached records in 2008). An equal number was observed in 2007. By comparison, 13 percent of cases investigated between 2004 and 2006 involved wireless networks. Incidentally, all were legitimate corporate WLANs rather than rogue devices deployed without proper authorization. There are many reasons for this decline, but better out-of-the-box security, wider use of encryption, and the necessity for proximity in order to conduct an attack are just a few. Web-based applications and remote access tools are, by their very operational nature, much more visible and accessible to external entities seeking a way into corporate networks.

A much smaller percentage of hacks targeted routers, switches, and other network devices. 2008 saw only a single instance in which a wireless network was exploited across our entire caseload.

Malware

ICSA Labs, an independent division of Verizon Business, provides credible third-party product testing and certification within the information security industry. When a member of the IR team discovers malicious software, or malware, during an investigation, it is sent to ICSA Labs for analysis. Investigators can then use this analysis to better help the customer with containment, removal, and recovery. The information that follows is based on this collaborative research.

During 2008, malware was involved in over one-third of the cases investigated and contributed to nine out of 10 of all records breached. In years past, malware was generally delivered in the form of self-replicating email viruses and network worms. The primary goal was rapid and widespread propagation, typically resulting in availability losses and extensive clean-up. In the last five years, these goals have shifted. Malware is now an essential component to nearly all large-scale data breach scenarios. Hacking gets the criminal in the door, but malware gets him the data. Naturally, the criminal will then want to minimize the chance of detection in order to maximize the amount of data stolen. For this reason, malware becomes ever more directed, innovative, and stealthy.

By a wide margin, the most common malware delivery method was the scenario in which an attacker compromised a system and then installed malware on it remotely. Perhaps more importantly, this delivery method accounts for 89 percent of the records breached in 2008. Seven infections occurred via websites, and, of these, four were "drive-by" downloads requiring no user interaction. The other three were explicitly downloaded and installed by employees. Only four cases in 2008 involved malware that exploited a patchable vulnerability. In all of these instances, the necessary patches to prevent infection were older than one year.

An important corollary to the infection vector is what the malware does once it is placed within the victim environment. In our previous report, we asserted that most malware captures and stores data locally, captures and sends data to a remote entity, or enables remote access to or control of the infected system. We also claimed that the ratio between these three functions was roughly equal. As seen in Figure 18, malware observed in 2008 exhibits a similar result.

Figure 17. Malware infection vector by number of breaches

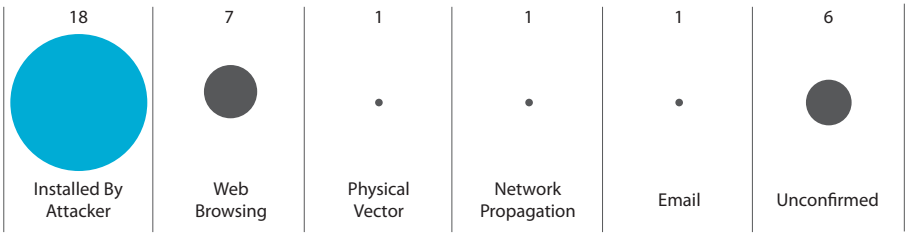


Figure 18. Malware functionality by number of breaches

The most prevalent of these functions were keyboard loggers or spyware. Typically, these are utilized to capture authentication credentials which are almost always sent to a remote attacker rather than stored locally for later retrieval. This is probably due to the small packet size that has a better chance of undetected egress. Criminals often use these credentials for subsequent and/or expanded attacks against corporate systems.

Responsible for 82 percent of total breached records in 2008, the most effective type of malware in terms of harvesting massive amounts of sensitive data is the "capture and store" variety. Attackers typically prefer this functionality for breaching payment card data and personally identifiable information (PII) since frequent exports of huge files containing millions of records is not the stealthiest of tactics. Of course, storing the payload on the victim's systems introduces its own challenges—namely, how to retrieve it. To solve this problem, the attacker will typically open up a backdoor in order to return to the system undetected over the (ordinarily) months that pass before the jig is up. Among our 2008 cases, investigators found backdoor or command shell tools in every instance where malware was capturing data to a local file.

The evolving use and functionality of malware in modern data compromise scenarios stems from the cybercrime market pressures described earlier in this report, but is also a direct response to the widespread adoption of various compliance standards and requirements. Organizations are implementing prescribed control measures in the manner and extent required to achieve and maintain compliance. Overall, this is a good thing for data protection within these organizations as common points of failure are (slowly but surely) being addressed. For instance, organizations are beginning to store less sensitive data as a part of normal business operations and encrypt what data they do retain. Less unencrypted information is flowing over public and private networks. Unfortunately, the criminals are not sitting around sulking about lost opportunities and dead-end business models; they are adapting.

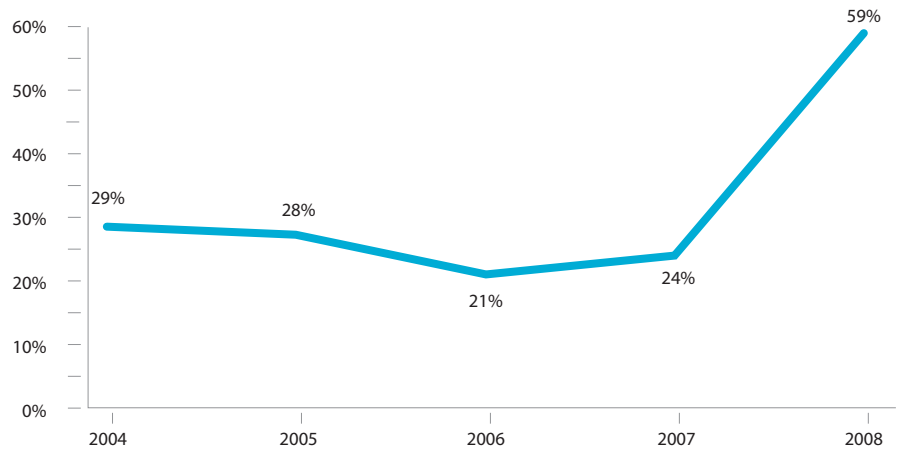
As organizations move to meet regulatory requirements, we have observed a manifest increase in attacks designed to circumvent certain controls implemented as part of that compliance process. Newer, more elaborate varieties of malware utilities bypass existing data controls and encryption, effectively creating vulnerable data stores that can later be retrieved from the victim environment. Examples of this include the usage of memory scrapers, sophisticated packet capture utilities, and malware that can identify and collect specific data sequences within unallocated disk space and from the pagefile.

Unfortunately, the criminals are not sitting around sulking about lost opportunities and dead-end business models; they are adapting.

Traditionally the term "stored data" has referred to nontransient items (i.e., in a log file or within a database on a hard drive, CD, or backup tape). However, the transient storage of information within a system's RAM is not typically discussed. Most application vendors do not encrypt data in memory and for years have considered RAM to be safe. With the advent of malware capable of parsing a system's RAM for sensitive information in real-time, however, this has become a soft-spot in the data security armor.

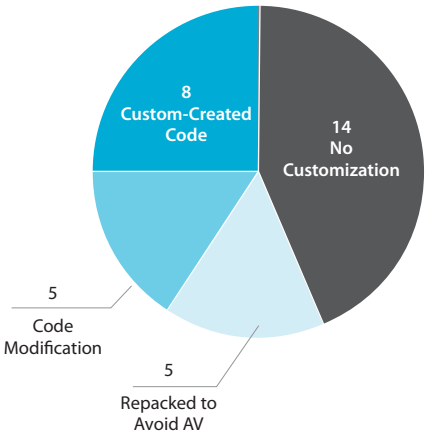
This expanded functionality, of course, doesn't happen magically. It requires authoring new malicious programs or significant modification of existing ones. This, in turn, requires considerable amounts of time, money, and expertise—not an investment many are willing or prepared to make. However, backed by the plenteous resources of organized crime and driven by the prospect of large hauls of valuable data, it is getting done. This trend was apparent in 2008, during which the percentage of customized malware more than doubled to 59 percent of all samples encountered.

Figure 19. Malware customization by percent of breaches involving malware



Among these cases, the amount and type of customization varied. A very general representation of this is given in Figure 20. Some attackers simply repacked existing malware so as to make its signature undetectable by antivirus software (AV) scanners. Others leveraged existing malicious code, but modified it for additional functionality or tailored it to the victim's environment. Most common in 2008, however, was malware that had (apparently) been created for the attack(s) entirely from scratch. In a rather sobering statistic, 85 percent of the 285 million records breached in the year were harvested by custom-created malware. It is possible that the code preexisted yet went unrecognized by the experts and tools at ICSA Labs, but this matters little to the overall point.

Figure 20. Malware customization by number of breaches



More to the point is that, besides being more capable and better adapted, most malware used for the purpose of compromising data is not detectable by modern AV. Unfortunately, many organizations rely on AV as the primary means of malware prevention and detection. AV is certainly a foundational control, but the continuing evolution of malware leaves security programs built solely upon AV for combating malware unstable at best.

Misuse and Abuse

Misuse refers to the use of organizational resources and/or privileges for any other purpose than for that which they were originally intended. For this reason, the category is particular to insiders and partners, as they are trusted by the organization to some degree.

Overall, 22 percent of breaches were caused by some form of misuse (see Figure 13). We find the fact that insider and partner misuse only accounted for 2 percent of all records compromised in 2008 to be counterintuitive since breaches involving the abuse of system privileges are usually quite damaging. 2008 was a rather unusual year, however, and this finding squares

Newer, more elaborate varieties of malware utilities bypass existing data controls and encryption, effectively creating vulnerable data stores that can later be retrieved from the victim environment.

perfectly with the results discussed earlier with respect to internal breaches. Furthermore, such attacks tend to be more narrowly focused (one only steals what one intends to use) and target data types that are not suited to the number of records measurement. One "record" of IP can be quite damaging but drowned out by millions of payment card numbers.

Table 3 shows the types of misuse observed among these cases. Not surprisingly, the abuse of system access and privileges was common. Most of these are committed by insiders with administrative privileges and are deliberate and malicious in nature. Nonmalicious policy violations, which are undoubtedly far more common for the whole of security incidents than these breach-specific numbers reflect, do contribute to data loss events but to a lesser extent. Such activity is very often a vector for the introduction of malware into the organization. Also observed were two instances of embezzlement.

The data in Table 4 shows that when employees engage in misuse, they tend to target larger data repositories. The majority of server-related breaches resulted from privilege abuse, while incidents involving workstations and laptops were associated with policy violations. It is interesting to note that whereas other studies have found portable media to be the leading cause of data breaches, we observed only a single instance in which such devices were used. Furthermore, in this particular case, the success of the breach did not hinge on its use; the USB media was merely a convenient method of moving data (which in our determination would have occurred anyway using other means if it weren't available).

Deceit and Social Attacks

This category encompasses the use of deception or misrepresentation to exploit people, security measures, procedures, or anything else that furthers the goal of data compromise. These actions can be conducted through both technical and nontechnical means. Common examples of deceit include social engineering and phishing scams, both of which we observed in the 2008 data set. Deceit was apparent in only 12 percent of our cases and these actions resulted in the compromise of 6 percent of all records.

Table 3. Types of misuse by number of breaches

Abuse of system access/privileges	15
Violation of other security policies	6
Violation of PC/email/web use policies	5
Embezzlement	2

Table 4. Types of assets misused by number of breaches

Database server	6
Application server	5
Laptop	5
File server	3
Public kiosk system	2
POS system	2
Workstation	2
Portable media	1

Table 5. Vectors of deceit and social attacks by number of breaches

Email	5
In-person	4
Web/Internet	3
Phone	1
Media or paper	1
Instant messaging	0

Table 6. Targets of deceit and social attacks by number of breaches

End-user	8
Partners or customers	2
Human resources	1
IT Admin. or Super-User	0
Helpdesk	0
Senior management	0

Table 7. Types of physical attacks by number of breaches

Theft of asset	5
System access (via keyboard)	2
Equipment or system tampering	2
Loss or misplacement of asset	0
Wiretapping	0
Observation or shoulder surfing	0
Assault or threat of harm	0

The majority of deceitful activity was carried out through email and/or the Internet. However, criminals have not become entirely impersonal; several instances of social engineering involved face-to-face and over-the-phone interaction. Web and email served as the medium of phishing and other scam-type attacks.

End-users proved to be the primary target of attacks employing deceit which, along with other findings in this report, call for more effective security awareness programs at the end-user level.

Physical Attacks

Data breaches resulting from physical attacks continue to rank low on our list, a finding that some will perceive to be out of sync with public sources of breach statistics. There are several justifiable reasons for this, all of which relate to our caseload. First, many physical incidents such as lost or stolen laptops never result in data compromise yet certain regulations require that the incident be publicly disclosed. Our data set, on the other hand, is comprised only of incidents in which an actual breach occurred. This has a very dramatic effect on results. Second, the nature of physical events often precludes the need for third-party investigation since there is less forensic evidence available for examination.

Only eight breaches in 2008 involved a physical attack. These are broken down in Table 7 (the values sum to nine because one case involved two distinct actions). Instances of theft involved workstations, documents, and backup tapes (no laptops) while system access was associated with online data repositories. Both cases of tampering dealt with PIN entry devices attached to POS systems in retail stores.

In addition to the type of physical attack, investigators noted where these breaches took place. Here again, the values exceed the number of breaches. This was due to one fairly elaborate scheme that unfolded over several venues. The main observation is that the ratio of external locations to internal locations is roughly equal.

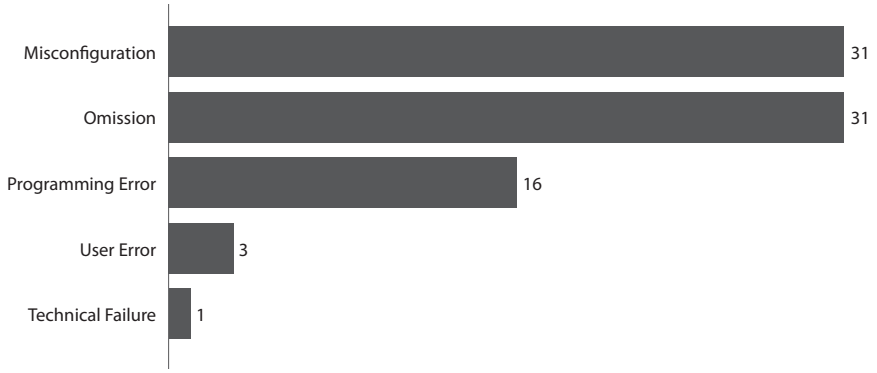
Table 8. Locations of physical attacks by number of breaches

Victim location: indoor, publicly accessible area (i.e., lobby, loading dock)	3
External location: store, restaurant, etc.	2
External location: employee home or car	2
External location: airport, train, subway, etc.	1
External location: business partner facilities	1
Victim location: indoor workspace (i.e., cubes, offices)	1
Victim location: indoor high security area (i.e., server room, R&D labs)	1
Victim location: outdoor, corporate grounds (i.e., parking lot)	0

Errors and Omissions

In its broadest sense, error is a contributing factor in nearly all data breaches. Poor decisions, misconfigurations, omissions, noncompliance, process breakdowns, and the like undoubtedly occur somewhere in the chain of events leading to the incident. Because error is so incredibly prevalent, only those errors which directly caused or significantly contributed to the compromise are considered by investigators for inclusion in this report. This distinction between error as a "direct cause" (or threat) vs. error as a "contributing factor" (failure in protection) is the reason behind the solid (cause) and shaded (factor) portions of the error bar in Figure 13. Intuitively, error is far less often the direct cause of a data breaches than it is one of several factors leading to its occurrence. In addition to causal and contributory, we further divide error into the types shown in Figure 21, which provides a relative distribution of each across 2008 cases in which an error was noted.

Figure 21. Types of error by number of breaches



Our use of *omission* and *misconfiguration* likely warrants some clarification. If an organization enacts a policy or procedure and then fails to follow through, we consider it an omission. In this sense, omission is not just another way to say the organization lacks adequate security measures (which is not considered an *error* in our approach) and is distinct from misconfiguration, which is a more active form of error during deployment and routine administration of systems. Whereas omissions dominated the 2004 to 2007 results, the scales are more balanced in 2008. We'd like to attribute this to dramatically improved assurance mechanisms over the last year, but—unfortunately—it is due to a methodological change on our part. Because the omissions category was so overpowering, we crafted a more strict definition and also added a new category for programming errors.

Poor decisions, misconfigurations, omissions, noncompliance, process breakdowns, and the like undoubtedly occur somewhere in the chain of events leading to the incident.

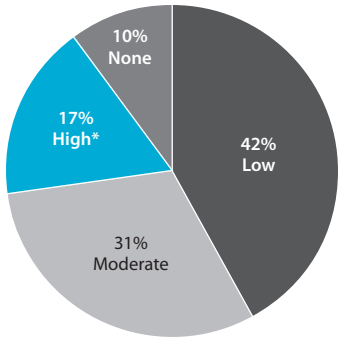
As Figure 21 illustrates, misconfiguration was the leading category of error contributing to data compromise, edging out omissions by a slight margin. 2008 saw quite a few instances of misconfigured or improperly constrained ACLs on perimeter devices. The use of default passwords despite policies forbidding them was a common example of omission. Clearly, procedures designed to identify and rectify these common mistakes would go a long way toward shoring up security efforts in most organizations. 21 percent of recorded errors stemmed from poor coding practices. Not surprisingly, errors of this variety are strongly related to the occurrence of SQL injection and other similar attacks. Writing secure code is not easy, but training and tools do exist, and—based on these results—developers would do well to avail themselves of both.

Attack Difficulty

The relative difficulty of attacks leading to data compromise is not only an excellent indicator of the current threat environment, but also the state of modern security programs. During each case, our investigators assess the details of the attack and classify it according to the following difficulty levels:

- **None:** No special skills or resources required. The average user could have done it.
- **Low:** Basic methods, no customization, and/or low resources required. Automated tools and script kiddies.

Figure 22. Attack difficulty by percent of breaches



*Highly difficult attacks accounted for 95% of all compromised records.

Comments or questions on this section?

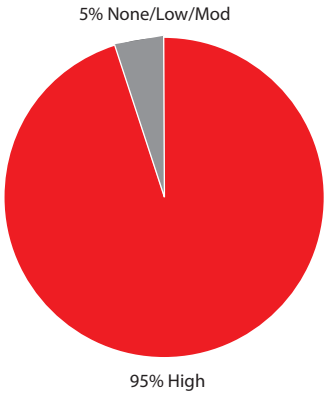
Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Attack Difficulty" post.

- **Moderate:** Skilled techniques, some customization, and/or significant resources required.
- **High:** Advanced skills, significant customization, and/or extensive resources required.

Although the rating admittedly involves a dose of subjectivity, we believe it to be a worthwhile metric.

One of the telltale findings from the previous DBIR was that more than half of breaches were caused by rather unsophisticated attacks. Within a few percentage points, our 2008 cases reveal a very similar statistic. The proportion of breaches requiring no special skills or resources (*None*) rose while low difficulty attacks fell to offset that gain. Attacks of moderate difficulty increased a few points and the percentage of highly difficult attacks remained the same. These subtle shifts are likely inconsequential and give little reason to believe that attack paradigms changed in 2008. On the whole, it would appear that criminals are still not required to work very hard to breach corporate information systems and 2008 was just a year like any other. Appearances, however, can be deceiving.

Figure 23. Attack difficulty by percent of records



While it may be true that the majority of breaches are not the result of highly skillful attacks, an alternate view of the data suggests that the really high-value targets require extensive effort. As Figure 23 plainly and powerfully demonstrates, these relatively few highly difficult attacks compromised 95 percent of the 285 million records across our caseload—a truly stunning statistic and one that is part of a larger story.

As discussed in the section detailing hacking activity, the techniques used by criminals to infiltrate corporate systems remain relatively low in sophistication. Exploiting default credentials and misconfigured ACLs does not require a great deal of skill but can be as effective a means of entry as more elaborate attacks. Increasingly, the complexity of an attack lies in the challenge of capturing and compromising the data and perpetuating access to systems. Since malware is typically the means by which this is accomplished, it has become the difficult component of modern attacks. Difficult attacks, therefore, are not necessarily difficult to prevent. It may sound trite, but the message here is that attacks are most efficiently and effectively prevented earlier rather than later.

In the 2008 DBIR, we stated that given enough time, resources and inclination, criminals can breach virtually any single organization they choose but do not have adequate resources to breach all organizations. Therefore, unless the value of the information to the criminal is inordinately high, it is not optimal for him to expend his limited resources on a hardened target while a softer one is available. These statistics make it clear that skilled cybercriminals are foregoing easy pickings in favor of inordinately high-value targets. Unfortunately, in 2008, they reaped the fruits of their labors.

Attack Targeting

Standard convention in the security industry classifies attacks into two broad categories: opportunistic and targeted. Due to significant grey area in this distinction, we find it useful to separate opportunistic attacks into two subgroups. The definitions are provided below:

- **Random Opportunistic:** Attacker(s) identified the victim while searching randomly or widely for weaknesses (i.e., scanning large address spaces) and then exploited the weakness.
- **Directed Opportunistic:** Although the victim was specifically selected, it was because they were known to have a particular weakness that the attacker(s) could exploit.
- **Fully Targeted:** The victim was first chosen as the target and then the attacker(s) determined a way to exploit them.

Figure 24. Targeted vs. opportunistic attacks by percent of breaches

*Targeted attacks accounted for 90% of all compromised records.

Given the message in the previous section that criminals seem to be selecting high-value targets and constructing elaborate attacks designed to breach their defenses, one might infer that targeted attacks rose in 2008. Based on our case evidence, that would be a valid inference. In fact, targeted attacks are at a five-year high and accounted for 90 percent of the total records compromised. In a comparison that speaks volumes about the uniqueness of 2008 in this regard, only 14 percent of the roughly 230 million records exposed among breaches we investigated from 2004 to 2007 were the result of fully targeted attacks. Financial services organizations are often singled out by criminals due to the large amounts of consumer data they process, transmit, and store.

As discussed in last year's report, we encounter many breaches that seem neither truly random nor fully targeted—particularly in the retail and food and beverage industries. In a very common example, the attacker exploits *Software X* at *Brand A* Stores and later learns that *Brand B* Stores also runs *Software X*. An attack is then directed at *Brand B* Stores but only because of a known exploitable weakness.

Compared to previous years, fewer breaches in 2008 resulted from purely random searching or scanning for exploitable weaknesses. One of the fundamental self-assessments every organization should undertake is to determine whether they are a Target of Choice or Target of Opportunity. If the former, expect and prepare for determined and sophisticated attacks. If the latter, minimize the opportunities presented so as to become less of a beacon for attack. At the very least, make sure your beacon shines less brightly than everyone else's.

One of the fundamental self-assessments every organization should undertake is to determine whether they are a Target of Choice or Target of Opportunity.

Comments or questions on this section?

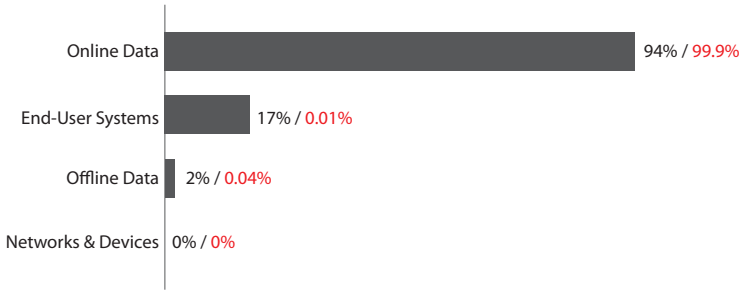
Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Attack Targeting" post.

Compromised Assets

We now turn attention to the types of information assets compromised during these attacks. The focus here is specifically on the endpoint rather than the pathway of the attack. To allow for comparison with the original report, the same asset groups are used in Figure 25. A more specific classification is provided in Table 7 for those interested in further details on these assets.

By a large margin and for the fifth year in a row, online data (consisting of various types of servers and applications) is the most frequently compromised asset. Furthermore, this group accounts for nearly all (99.9 percent) of the 285 million records breached across our 2008 caseload. As seen in Figure 25, this percentage represents a high mark for the five-year period of this study. Clearly, large and remotely accessible stores of data remain the target of cybercriminal activity.

Figure 25. Asset classes by percent of breaches (black) and records (red)



Looking a bit closer at online data in Table 9, POS systems were most frequently compromised but accounted for only a small portion of total records. Intuitively, these breaches predominantly afflicted the retail and the food and beverage industries. Databases rank second in terms of caseload but yielded the majority of breached data. All together, other types of online data listed in Table 9 factored into a third of breaches, but—of those—only the application servers had substantial losses of data.

One final point of interest concerning online data is that of virtualization. After the release of last year's report, there was some discussion on what our results meant for virtualization. As such, we made a point to note whether assets involved in a case were virtualized. We encountered five instances among 2008 cases but none were believed to have contributed to the breach in any way.

Although much angst and security funding is given to offline data, mobile devices, and end-user systems, these assets are simply not a major point of compromise within the data set available to us for examination. It is indisputable that employees misuse portable media and laptops go missing, and—based on public breach disclosure lists like DatalossDB.org* and ID Theft Center**—it is also evident that large numbers of records are reported exposed from related incidents. That such trends are not reflected in the results above is a by-product of our caseload, this data set, and the general nature of data compromise. Verizon Business is not often engaged to investigate lost devices and this data set is culled down to only cases in which a breach was confirmed. Furthermore, information on lost or stolen laptops and media is rarely accessed and abused by criminals; data-at-risk—though often necessary to report—is not the same as actual data compromise.

*http//datalossdb.org

**http://www.idtheftcenter.org

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Compromised Assets" post.

Figure 26. Percent of records breached from online data assets

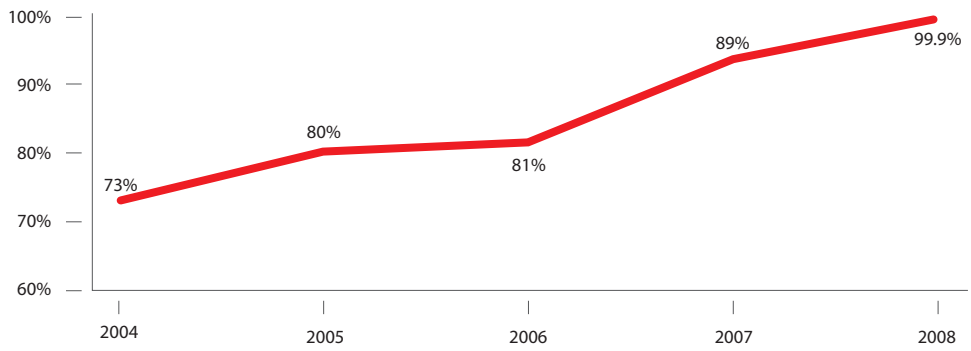


Table 9. Detailed listing of compromised assets by percentage of breaches and records

Asset	Asset Group	% of Breaches	% of Records
POS system	Online Data	32%	6%
Database server	Online Data	30%	75%
Application server	Online Data	12%	19%
Web server	Online Data	10%	0.004%
File server	Online Data	8%	0.1%
Public kiosk system	Online Data	2%	0.4%
Authentication / Directory server	Online Data	2%	0.1%
Backup tapes	Offline Data	1%	0.04%
Documents	Offline Data	1%	0.000%
Workstation	End-User System	8%	0.01%
Laptop	End-User System	4%	0.000%
PIN Entry Device	End-User System	2%	0.004%

Compromised Data

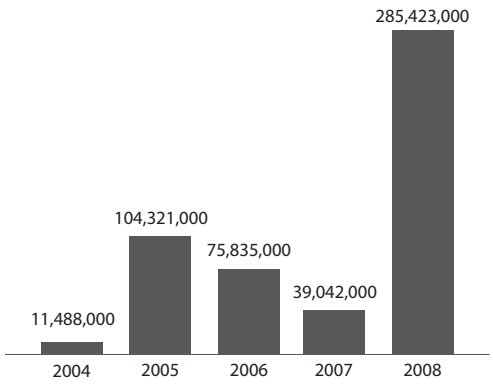
2008 was a milestone year in terms of the sheer number of records compromised in breaches investigated by Verizon Business. More records were breached in 2008 than any other single year, and—astoundingly—more than in the previous four years combined. In the last five years, our team has investigated cases totaling more than half a billion records.

That 2008 is such an anomaly in this regard is the result of a few very large breaches. The top five breaches account for 93 percent of total records compromised. The mean number of records per breach was approximately 4.5 million, while the median was 37,847. From these statistics, the data set is obviously skewed right, which can also be seen in the distribution in Figure 28.

As we pointed out last year, one of the more critical components of the investigative process is determining the types of data compromised as a result of the breach. This determination largely decides the organization's response strategy. For instance, certain types of data require public disclosure and/or notification. Some are highly regulated while others require monitoring for fraudulent activities. In addition to the victim organization, millions of individuals could potentially be affected by the compromise, and they must be informed as well.

Overall, the results for 2008 are fairly consistent with those of the previous four years. As a percentage of caseload, payment card breaches remain near the 80 percent mark and far outnumber other data types. They consume 98 percent of all records compromised in the year. While other types of data are sought by certain groups (i.e., competitors may target IP), the vast majority of cybercriminals are looking for a quick and easy payoff. Payment cards certainly fit the bill. As a testimony to this, fraudulent use of stolen card data was confirmed in 83 percent of our cases.

Figure 27. Number of records compromised per year in breaches investigated by Verizon Business

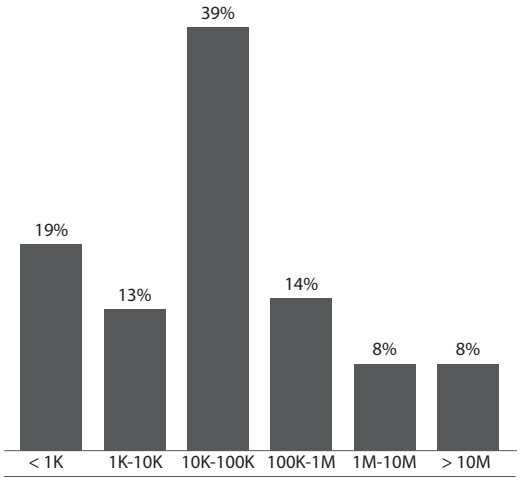


The payment card data percentage above includes not only the typical data sequences sufficient for card-present and card-not-present fraud but also PIN information associated with consumer payment card accounts. The latter was increasingly targeted in 2008. These attacks involve the identification and compromise of stores of magnetic-stripe data, together with PIN, setting the stage for more damaging forms of identity fraud. This includes, for example, counterfeit ATM withdrawals and other PIN-based transactions, often leading to actual cash and related assets being stolen directly from the consumer. While statistically not a large percentage of our overall caseload in 2008, attacks

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Compromised Data" post.

Figure 28. Distribution of breach size by number of records

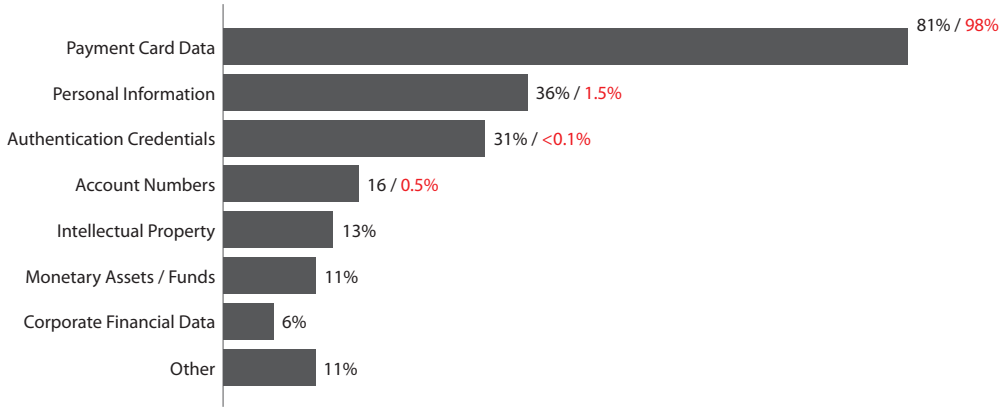


against PIN information represent individual data theft cases having the largest aggregate exposure in terms of unique records. In other words, PIN-based attacks and many of the very large compromises from the past year go hand in hand.

PII was the second-most compromised data type, which only stands to reason as it is also useful in fraudulent activity. As the name implies, PII includes elements of a person's unique identity such as name, identification number, etc. At 30 percent of cases, stolen authentication credentials in 2008 doubled the 2004 to 2007 figure. Whether this is indicative of some change in criminal strategy remains unknown (it had been declining each year prior to 2008). Credentials certainly give attackers the prospect of increased access for illicit activity and based on the results from the hacking section earlier in this report, they appear to be exploiting that advantage.

Though still a smaller share of breaches, theft of intellectual property within our caseload rose to a five-year high in 2008. Such incidents, while infrequent, can be business-changing events. Account numbers and monetary assets were added to the list of data types in 2008. Neither was extremely common but multiple instances of each were observed. Account numbers (often bank accounts) are distinct from payment card numbers. For obvious reasons, these two data types often appeared in concert; millions of dollars were stolen directly from compromised accounts in 2008.

Figure 29. Compromised data types by percent of breaches (black) and records (red)*



*Due to the nature of certain data, not all types listed in Figure 29 are included in percentage of records statistics.

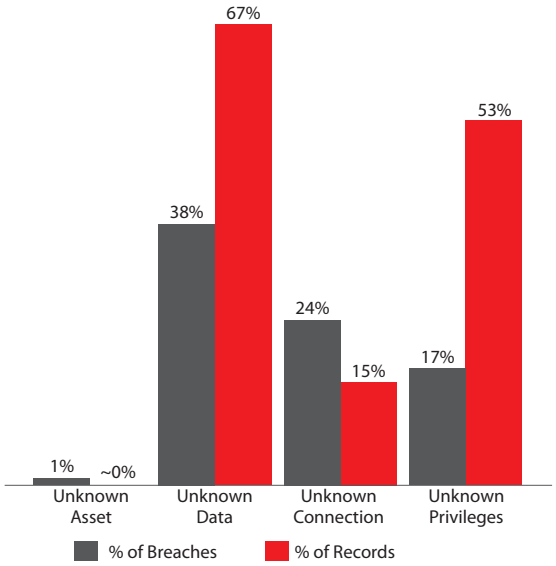
Unknown Unknowns

Several years ago, a strong pattern emerged among cases wherein the investigation inevitably led to unknown, overlooked, or forgotten assets. At some point, we took to referring to these frequent scenarios as "unknown unknowns," and they include any of the following:

- A **system** unknown to the organization (or business group affected)
- A system storing **data** that the organization did not know existed on that system
- A system that had unknown network **connections** or accessibility
- A system that had unknown accounts or **privileges**

Figure 30 shows the percentage of 2008 cases in which each of these contributed to a data compromise.

Figure 30. Unknown unknowns by percent of breaches



Roughly half of all breaches investigated in 2008 revealed at least one type of unknown. Though substantial, this statistic is considerably lower than the 90 percent given for the previous four years. A completely unknown asset factored into only one case, while unknown connections remained near the 25 percent mark and unknown privileges rose by 7 percent. Far fewer breaches involved data that the victim did not know existed on the system and this fact accounted for most of the difference between the two data sets. Rather than some kind of sampling effect, there are several legitimate reasons for this change.

You may remember from the 2008 Data Breach Investigations Supplemental Report that financial services organizations were less afflicted with unknowns than other industries. Because these organizations represent a higher proportion of our caseload in 2008, it stands to reason that they would have a downward pull on this statistic. Incidentally, this may also explain the rise in unknown privileges and drop in unknown connections. Second, based on findings from other sections of this report, it is evident that we observed a higher-than-normal number of highly targeted and customized attacks this year. The criminals appear to be going for the "crown jewels," and the victim typically knows where those reside. Rather than the result of unintended data migration (though this certainly occurs), more of these breaches stem from a combination of the criminal's resolve and the victim's failure to adequately protect systems they know to contain sensitive data. Still, many did involve unknown data of a sort—but it was data captured by malware and stored on the system unbeknownst to the organization.

Comments or questions on this section?

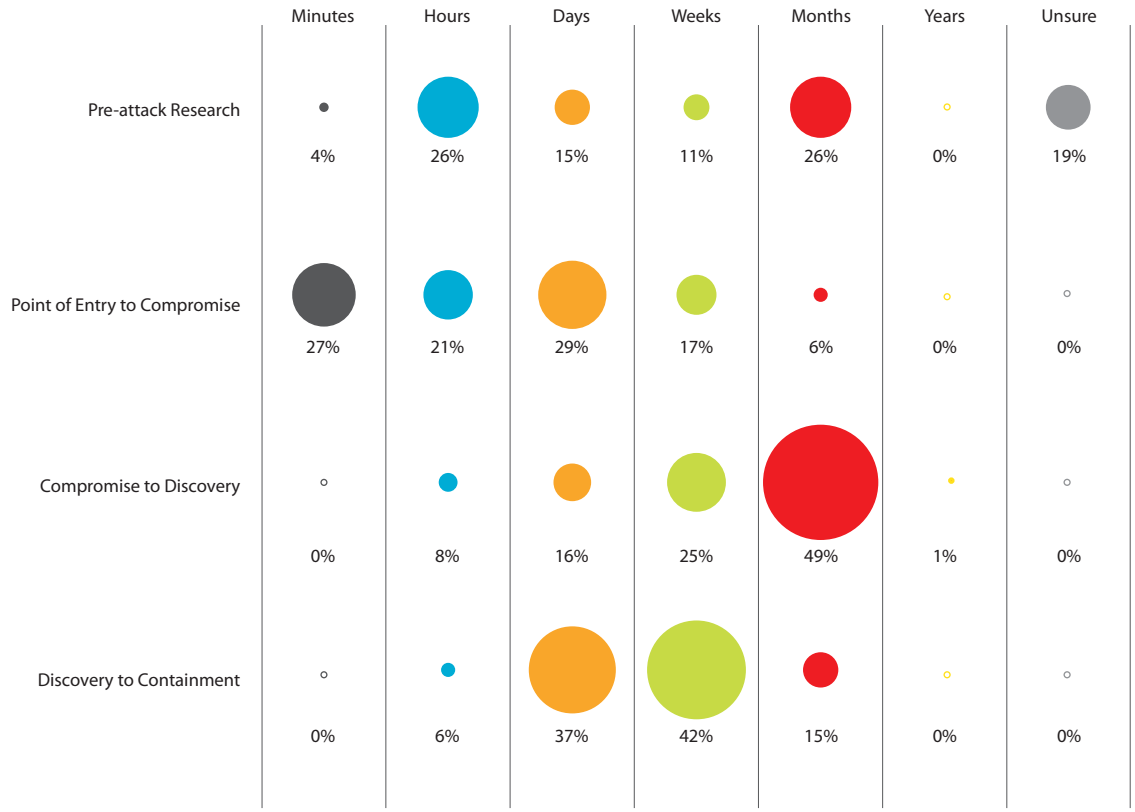
Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Unknown Unknowns" post.

Though breaches involving unknown data have seen decline, poor account and privilege management has become more problematic in recent years both as a percentage of cases and in the number of compromised records. Additionally, Figure 30 reminds us that the consequences of the "unknown unknown" remains high. This suggests that increasing visibility and reducing variability in the IT operating environment should be a top priority of risk management efforts.

Time Span of Breach Events

The timeline of events leading up to and following a data breach varies greatly depending on a multitude of factors. Some attacks unfold rapidly while others require lengthy planning and execution. For the purposes of analysis, we separate an incident into three major phases: point of entry to compromise, compromise to discovery, and discovery to containment. Information regarding the amount of time spent on pre-attack research is also of interest and recorded when obtainable. The figures below depict the span of time elapsed during each phase.

Figure 31. Time span of breach events by percent of breaches



Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Time Span of Breach Events" post.

Pre-Attack Research

Though solid evidence of pre-attack activity is often elusive, investigators attempt to discern as much as possible about an attack prior to the point of entry. Doing so often uncovers hints as to how the target was acquired, what reconnaissance methods were employed, and—most importantly—what signs might have forewarned the victim of an impending attack.

In just under half of our cases in 2008, investigators found at least some indication of pre-attack research. This most often involved basic system footprinting, scanning, and enumeration, but instances in which the attacker scouted the victim's premises were observed as well. The latter were designed to uncover the make and model of POS systems or study traffic patterns to identify the optimal time for tampering with the PIN Entry Device (PED). In several cases, the assailant was able to obtain private information (a vendor's customer lists, for instance) and leveraged it to further customize or to broaden an attack. Per Figure 31, the time required to conduct these activities varied from minutes to months.

Point of Entry to Compromise

After breaching the perimeter, intruders typically explore the victim's network and systems until finding their desired plunder. The length of time necessary to accomplish this largely depends on factors such as the attacker's prior knowledge, skill set, familiarity with the environment, the method of attack, and the strength of the victim's defenses. Overall, 2008 cases follow a very similar distribution to those in our 2004 to 2007 data set. Though more compromises occurred within minutes, fewer required hours of effort. Criminals can still access the data in a matter of minutes or hours about half of the time. The proportion corresponding to days, weeks, months, and years remained nearly identical.

Compromise to Discovery

On the whole, organizations discovered breaches slightly quicker in 2008. However, lest we confuse "quicker" with "quickly," this statement needs some additional context. Breaches still go undiscovered and uncontained for weeks or months in 75 percent of cases. It is doubtful that any chief security officer anywhere would call this "quick." Nevertheless, rather than the supermajority observed in previous years, 2008 investigations revealed that only a slight majority of breaches took months or years to discover. On the positive side, we saw comparatively more weeks instead of months and the number detected within hours increased by 5 percent. We can only hope such trends will continue.

On the whole, organizations discovered breaches slightly quicker in 2008. However, lest we confuse "quicker" with "quickly," this statement needs some additional context. Breaches still go undiscovered and uncontained for weeks or months in 75 percent of cases. It is doubtful that any chief security officer anywhere would call this "quick".

Discovery to Containment

In the previous report, we presented this phase under the heading "Discovery to Mitigation." Afterward, it was determined that "Containment" provided better clarity as to our meaning. By containment, we essentially mean to "stop the bleeding"; it does not refer to complete remediation of the root problem(s). Once a breach is contained, unauthorized access is cut-off and information is no longer exposed.

As one can imagine, containing a data compromise situation as quickly as possible is paramount to the response process. Unfortunately, this goal is rarely met. The majority of breaches required weeks or more to reign in and very few were contained in less than a couple days. We believe insufficient preparedness to be the crux of the issue—many organizations are simply not ready to respond when a crisis occurs. As a testimony to this, findings pertaining to response practices in place at the time of the incident show that a large proportion of organizations were not adequately prepared to handle an incident. Specific findings concerning the maturity of incident response practices are presented in the "Discovery and Response" section that follows.

Discovery and Response

In addition to time-to-discovery and time-to-containment statistics, information surrounding the method of discovery, detection capabilities, and incident response practices can help develop a clearer picture around what contributes to the frequency and severity of breaches. As this topic is so critical and yet the findings from our previous report so deplorable, new metrics were added for 2008 to allow further study. We begin with the method of discovery.

Discovery Methods

In last year's study, we reviewed how organizations became aware that a breach had occurred. 2008 investigations reveal no significant differences in the method of discovery; the vast majority of breaches continue to be discovered by a third party. Some will notice a divergence in Figure 32 with respect to the "Notification by Third Party" category presented in the previous report. Because this discovery method was and is so over-represented, we separated fraud detection from other types and reasons of third-party notification. From this view, it is clear that breaches go unnoticed until a sufficient number of fraudulent transactions enable a third party to isolate the point of compromise.

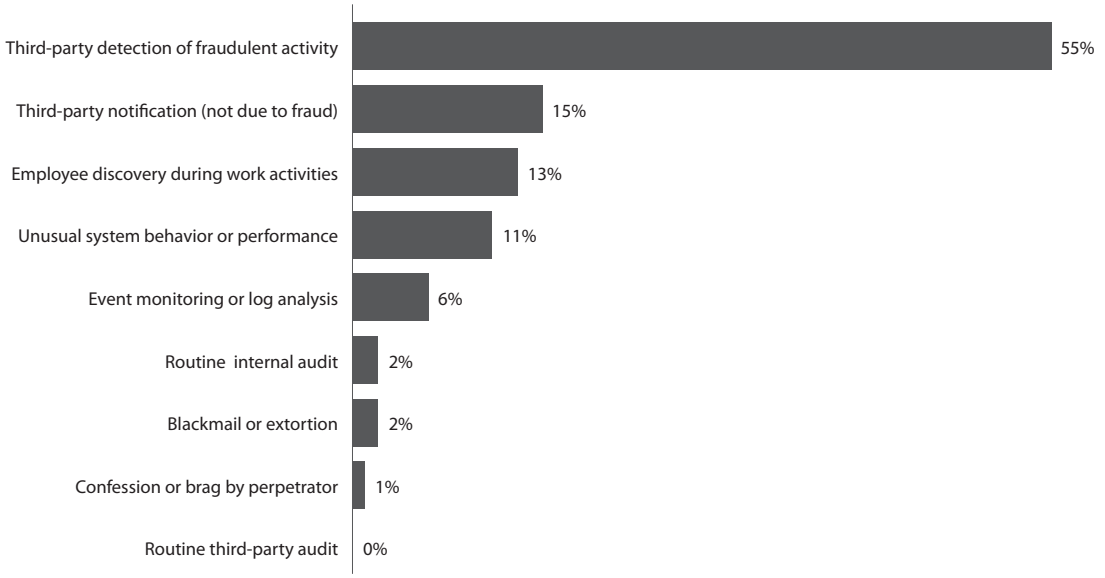
Employees happening upon breaches during their normal work activities remain a distant second (to the combined third-party detection or notification categories) at 13 percent, followed closely by unusual system behavior at 11 percent. These two methods, while unplanned, seem to be an organization's best chance of discovering a breach on their own. That's hard to believe, but it's also hard to argue with five years of data that all reveal a similar finding.

A simplified view of breach discovery methods is given in Figure 36, which concisely contrasts the regularity of third party and internal discovery. As with prior years, "active" measures (referring to those that are specifically designed for detection) taken by the organization detect only a small proportion of breaches. It is clear that breach discovery remains a largely passive endeavor in 2008.

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/,
and look for the "Discovery and Response" post.

Figure 32. Breach discovery methods by percent of breaches



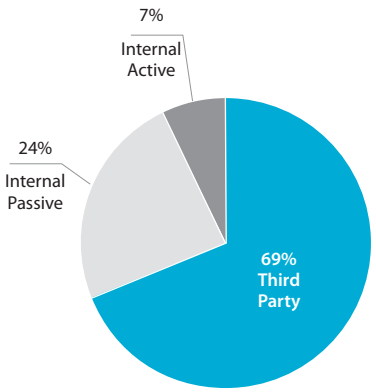
The apparent ineffectiveness of event monitoring and log analysis continues to be somewhat of an enigma. The opportunity for detection is there; investigators noted that 66 percent of victims had sufficient evidence available within their logs to discover the breach had they been more diligent in analyzing such resources. Though lower than in previous years (it was 82 percent from 2004 to 2007), this finding still suggests that realized effectiveness remains much lower than potential effectiveness. Of course, the extent of this underperformance depends greatly upon the deployment of these technologies. We examine this in the next section.

Utilization of Detective Controls

The information presented in the previous section raises obvious questions concerning the prevalence and operational maturity of detection-oriented controls. To help find answers, investigators gathered information on the utilization of these controls among organizations within our 2008 caseload. The findings are presented in Figure 34.

The data suggest that, beyond basic system and device logs, the utilization of detective controls among breach victims is relatively low. If your thought process is anything akin to ours, your eye went immediately to the 30 percent

Figure 33. Breach discovery methods, simplified



attributed to intrusion detection systems and you muttered "Gee—that seems low." The temptation is to write this off as a byproduct of a sample that is obviously comprised of small organizations and immature security programs. This is not the case and disregarding these results as non-reflective of larger enterprises would be a mistake. Though still unexpectedly low, there are several valid reasons that help explain this result.

First, a number of large organizations that one would fully expect to have intrusion detection systems (IDS) did not. Others had deployed IDS but had not activated them. Still more (about 12 percent) were monitoring network activity with IDS but not in the locations or assets involved in the breach. In other cases, we were simply unable to confirm the existence of IDS and other detective capabilities (for various reasons). Finally, it is worth noting that some organizations represented in the study did not have prescriptive compliance requirements for event detection.

Of the five breaches that were detected by event monitoring and log analysis, the IR team noted that three collected syslogs and ran IDSs, one collected syslogs and regularly reviewed them, and one employed all detective controls listed in Figure 38.

Figure 34. Detective controls by percent of breach victims

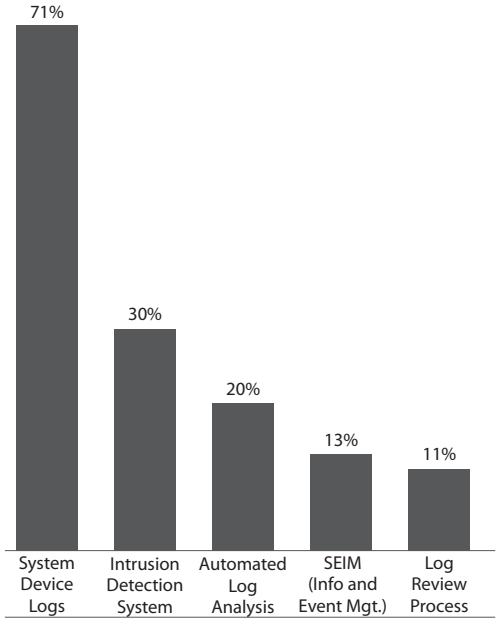


Figure 35. Incident response practices by percent of breach victims

Incident Response Practices

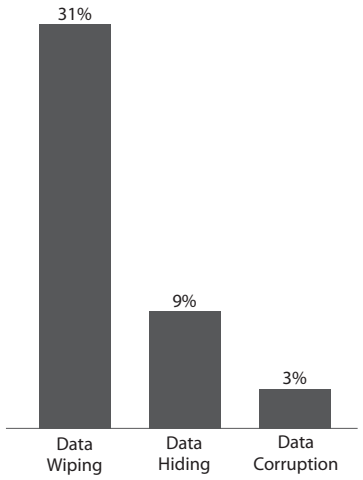
During 2008 cases, investigators also noted the existence of the Incident Response practices shown in Figure 35. Overall, the results suggest that, in addition to low detective capability, most organizations are ill-equipped to adequately respond to breaches when they are discovered. We found it especially surprising that only 28 percent of victims had an incident response plan in place.

Anti-Forensics

When criminals decide to break into corporate networks and steal data, getting caught is not typically part of their plan. Many will attempt to muddy the crime scene and remove evidence of their actions to foil post-incident investigations. Such activity is referred to as anti-forensics and is at least partially responsible for the breach discovery and response struggles discussed in the previous section.

Investigators found signs of anti-forensics in over one-third of cases in 2008. The most common forms of anti-forensics measures observed in the field are Data Wiping, Data Hiding, and Data Corruption. The prevalence of these among investigations conducted in the last year is shown in Figure 39. The pervasiveness of Data Wiping, which includes removal and deletion, comes as no surprise. Many commercial off-the-shelf products are available that perform this function (there are many legitimate uses for wiping data) as well as freeware utilities. With respect to Data Hiding, the use of steganography has remained relatively rare and flat year-over-year. On the other hand, the use of encryption for the purposes of Data Hiding has risen by almost 10 percent. Where Data Corruption was observed, it was mostly manifested as log tampering.

Figure 36. Types of anti-forensics encountered by percent of breaches



Intuitively, the breaches in which anti-forensics were utilized tended to be larger in size and longer in duration. For many of the smaller "smash and grab" attacks, destroying evidence is simply not worth the effort (plus those attacks are often perpetrated by attackers of lower skill who are ignorant of anti-forensic measures). In this sense, the use of anti-forensics is progressing beyond its roots as a defensive tactic used to avoid prosecution and is increasingly employed in an offensive capacity to expand and extend the attacker's ability to compromise data.

Comments or questions on this section?

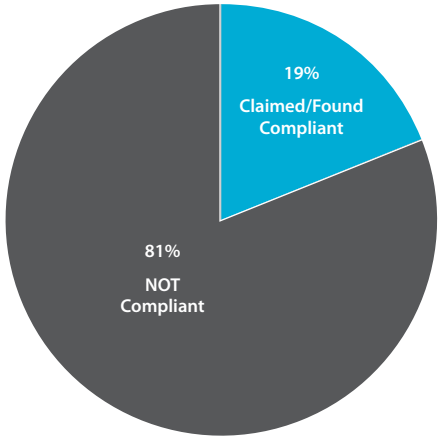
Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "Anti-Forensics" post.

Payment Card Industry Data Security Standard

A great deal of public discussion has taken place regarding the effectiveness of regulations and control guidelines to prevent breaches, but empirical study of the topic remains scarce. While it is not the purpose of this report to provide this much-needed analysis, our 2008 data set does offer some information concerning incidents where the organization is required to comply with the Payment Card Industry Data Security Standard (PCI DSS).

PCI DSS is a set of control requirements created to help organizations protect cardholder information. Based on the representation of financial institutions and retailers in our data set, it is no surprise that the large majority of organizations within our caseload are subject to the requirements set forth within PCI DSS. Because these were data compromise cases, an obvious question arises as to the compliance status of the breach victims. The answer to this question is given in Figure 40 but additional clarification is needed to ensure proper interpretation.

Figure 37. PCI compliance status based on last assessment by percent of breach victims



Over three-quarters of organizations suffering payment card breaches within our caseload were found not compliant with PCI DSS or had never been audited. This status was not determined by our Investigative Response team but rather by the victim's attestation or Qualified Security Assessor (QSA). Far more interesting than these non-compliant organizations, however, are the 19 percent that had been found compliant during their last assessment. Does the fact that these organizations were deemed compliant by their QSA and yet still suffered an incident mean that PCI DSS was ineffective at preventing those breaches? Our findings do not support such a conclusion.

Due to the point-in-time nature of PCI assessments, it is possible that an organization deemed compliant at its last audit may not still be compliant at the time of the breach. Furthermore, PCI compliance is not an absolute guarantee against breaches nor is the assessment process always consistent. Whatever the reason, post-breach reviews conducted by our IR team reveal some very interesting findings.

At the culmination of a forensic engagement, the investigator working the case performs a review of which PCI DSS requirements were and were not in place (and adequate) at the time of the breach. The results of this assessment are recorded

Comments or questions on this section?

Visit http://securityblog.verizonbusiness.com/category/2009dbir/, and look for the "PCI DSS" post.

in the PCI Requirements Matrix, appended to the case report, and then conveyed to the relevant payment card brands. This exercise is not an official PCI certification audit and it does nothing to uphold or overrule the victim's compliance status. It does, however, provide useful insight into which requirements tend to be deficient relative to organizations that suffer data breaches. The aggregate results of post-breach PCI reviews conducted during 2008 are tallied in Table 10.

Table10. Results of post-breach PCI DSS reviews conducted by Verizon Business IR. Values represent the percentage of organizations for which each requirement was found to be in place.

Build and Maintain a Secure Network	**Compliance**
Requirement 1: Install and maintain a firewall configuration to protect data.	30%
Requirement 2: Do not use vendor-supplied defaults for system passwords and other security parameters.	49%
Protect Cardholder Data	
Requirement 3: Protect stored data.	11%
Requirement 4: Encrypt transmission of cardholder data and sensitive information across public networks.	68%
Maintain a Vulnerability Management Program	
Requirement 5: Use and regularly update AV.	62%
Requirement 6: Develop and maintain secure systems and applications.	5%
Implement Strong Access Control Measures	
Requirement 7: Restrict access to data by business need-to-know.	24%
Requirement 8: Assign a unique ID to each person with computer access.	19%
Requirement 9: Restrict physical access to cardholder data.	43%
Regularly Monitor and Test Networks	
Requirement 10: Track and monitor all access to network resources and cardholder data.	5%
Requirement 11: Regularly test security systems and processes.	14%
Maintain an Information Security Policy	
Requirement 12: Maintain a policy that addresses information security.	14%

When reviewing the percentages for each requirement above, several very interesting statistics pop out. Requirements 3, 6, and 10—which many organizations complain are the most onerous—are indeed the least compliant across our caseload. When one considers the prevalence of unnecessary and/or unknown data stores, frequency of SQL injection attacks, and lengthy compromise-to-discovery periods discussed extensively in this (and our last) report, this finding is hardly surprising. This trio of deficiencies factored heavily into many of the largest breaches investigated by our team over the past five years. Unfortunately, these statistics—and those discussed earlier regarding the low utilization of discovery and detective controls—suggest that attacks exploiting these areas will continue to be a challenge for the foreseeable future.

In other words, the typical organization had met less than a third of the requirements in PCI DSS. Some fared much better, some much worse, but the point made by the data before us is this: these breaches, in general, did not occur in organizations that were highly compliant with PCI DSS.

On the other end of the spectrum, encrypting data over public networks and antivirus software exhibit fairly high levels of compliance (68 percent and 62 percent respectively). Unfortunately, they have little to do with common attack modalities. As discussed earlier in this report, information is either captured off private networks or within systems and AV software is of little effect against customized malware installed on a system under the control of a remote attacker. Nevertheless, antivirus software certainly plays a major role in enterprise defenses and the fact that 40 percent of victims were not using it (or not updating it) is quite surprising.

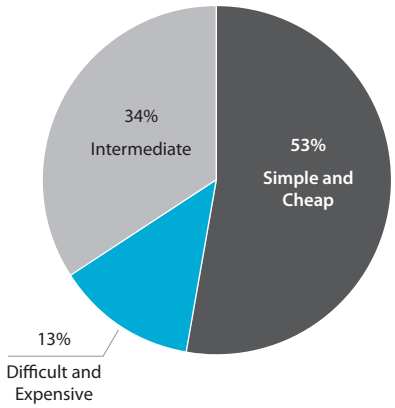
Incredibly, 51 percent of the victims were using vendor default passwords on systems that handle sensitive data. In over 80% of our cases, the victim organizations used shared accounts for system access and did not assign unique IDs to users. That the low adherence to these requirements played a role in the numerous breaches attributed to default and shared credentials used for third-party remote administration services is indisputable.

While the requirement-centric view of this information is interesting, the organization-centric view makes a compelling point as well. Our unofficial post-breach PCI reviews found the average compliance rate across victims in our caseload to be 29 percent of the 12 requirements. In other words, the typical organization had met less than a third of the requirements in PCI DSS. Some fared much better, some much worse but the point made by the data before us is this: these breaches, in general, did not occur in organizations that were highly compliant with PCI DSS.

Conclusions and Recommendations

The conclusion to our previous report began with the statistic that, in 87 percent of cases, investigators concluded that the breach could have been avoided if reasonable security controls had been in place at the time of the incident. Some felt our use of "reasonable" was not entirely clear (you can read our explanation here*), so we opted for a different route this time. That route, however, arrived at the very same conclusion—just worded differently.

Figure 38. Description of the effort and expense of recommended preventative measures by percent of breaches



In 2008, investigators concluded that 87 percent of breaches could have been avoided through the implementation of simple or intermediate controls. All of these were the standard, run-of-the-mill practices that we in the industry see and use everyday. Costly controls (in terms of effort and expense) were recommended as the most efficient and effective means of avoiding the breach in only 13 percent of cases. Furthermore, most of these—though costly—were standard security controls. Figure 39 provides a different classification of these same controls.

At the conclusion of the original DBIR, we made a number of recommendations based directly on findings from our 2004 through 2007 breach investigations. Additionally, we called for a shift in the data protection and incident response mentality that pervades the industry. This year, we would like to reiterate that call and recapitulate those recommendations as well as provide additional guidance based on our 2008 caseload.

An abbreviated version of recommendations from the 2008 DBIR follows. They are as relevant today as they were last year.

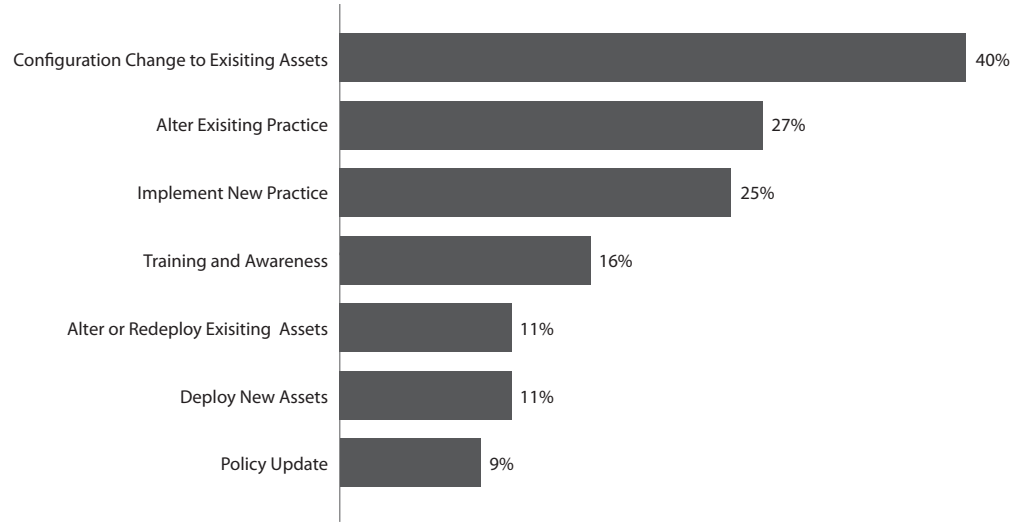
Align process with policy: Many organizations set security policies and procedures yet fail to implement them consistently. Controls focused on accountability and ensuring that policies are carried out can be extremely effective in mitigating the risk of a data breach.

Achieve essential, and then worry about excellent: We find that many organizations achieve very high levels of security in numerous areas but neglect others. Criminals will almost always prefer the easier route. Identifying a set of essential controls and ensuring their implementation across the organization without exception, and then moving on to more advanced controls where needed is a superior strategy against real-world attacks.

Secure business partner connections: Basic partner-facing security measures as well as security assessments, contractual agreements, and improved management of shared assets are all viewed as beneficial in managing partner-related risk.

*http://securityblog.verizonbusiness.com/2008/06/19/reasonable-controls/

Figure 39. Simple categorization of recommended preventive measures by percent of breaches



Create a data retention plan: Clearly, knowing what information is present within the organization, its purpose within the business model, where it flows, and where it resides is foundational to its protection. Where not necessitated by valid business needs, a strong effort should be made to minimize the retention and replication of data.

Control data with transaction zones: Based on data discovery and classification processes, organizations should separate different areas of risk into transaction zones. These zones allow for more comprehensive control implementations to include but not be limited to stronger access control, logging, monitoring, and alerting.

Monitor event logs: All too often, evidence of events leading to breaches was available to the victim but this information was neither noticed nor acted upon. Processes that provide sensible, efficient, and effective monitoring and response are critical to protecting data.

Create an Incident Response Plan: If and when a breach is suspected to have occurred, the victim organization must be ready to respond. An effective Incident Response Plan helps minimize the scale of a breach and ensures that evidence is collected in the proper manner.

Increase awareness: Delivered effectively, training that educates employees about the risks of data compromise, their role in prevention, and how to respond in the event of an incident can be an important line of defense and discovery.

Engage in mock incident testing: In order to operate efficiently, organizations should undergo routine IR training that covers response strategies, threat identification, threat classification, process definition, proper evidence handling, and mock scenarios.

In addition to these "oldies but goodies," we offer the following new or expanded recommendations based on our analysis of 2008 data. As before, these recommendations are in no way a comprehensive strategy for enterprise data protection and certainly cannot guarantee against compromise. Rather, they are a consequence of the experience gained through the hundreds of data breach investigations and are intended to highlight specific problem areas common to many organizations within our caseload.

Changing default credentials is key: More criminals breached corporate assets through default credentials than any other single method in 2008. This certainly falls under the prior "achieve essential, and then worry about excellent" recommendation, but the prevalence of this specific failure justifies its special mention. It is not as though the victims had no idea that changing usernames and passwords was important; they just neglected to do so on a few assets (and the intruders, of course, found those) or they assumed the third party responsible for managing them would do so. The lesson? Don't assume that your staff or your partner's consistently follows through on all policies and procedures.

Avoid shared credentials: Another obvious yet frequently omitted and oft-exploited problem. Along with changing default credentials, organizations should ensure that passwords are unique and not shared among users or used on different systems. The use of shared credentials allowed quite a few breaches in 2008 to grow from a single point of compromise to multiple incidents. This was especially problematic for assets managed by a third party.

User account review: This is related to the two prior recommendations but worth calling out individually. Default and shared credentials, a growing rate (and large number of records lost) of breaches involving unknown privileges, and years of experience lead us to believe in the value of reviewing user accounts on a regular basis. The review should consist of a formal process to confirm that active accounts are valid, necessary, properly configured, and given appropriate (preferably least) privileges.

Application testing and code review: SQL injection attacks, cross-site scripting, authentication bypass, and exploitation of session variables contributed to nearly half of breaches attributed to hacking or network intrusion. It is no secret that attackers are moving up the stack and targeting the application layer. Why don't our defenses follow suit? As with everything else, put out the fires first: even lightweight web application scanning and testing would have found most of the problems that led to major breaches in the past year. Next, include regular reviews of architecture, privileges, and source code. Incorporating a Security Development Life-Cycle (SDLC) approach for application development is recommended as well. Finally, help your developers learn to appreciate and write more secure code.

Smarter patch management strategies: For every vulnerability exploited by hacking and malware attacks in 2008, the patch necessary to prevent the breach had been available for at least six months prior to the incident. In fact, all but one had been around for a year or more. While it may seem logical to conclude that organizations aren't patching fast enough, this is not the correct interpretation. All of these organizations had patch cycles well below the six month mark. The

problem throughout all five years of this study has far more to do with scope than speed. Organizations would find much more value if they divert resources from patching ever-faster to patching more consistently and comprehensively.

Human resources termination procedures: Several breaches in the last year were the result of malicious activity on the part of a recently terminated (or notified) employee. Each organization should have a comprehensive employee termination policy that identifies parties responsible for the termination process, outlines procedures and checklists for termination, ensures the retrieval of all organizational property, and—most importantly—establishes a process for quickly disabling user accounts and removal of all access permissions.

Enable application logs and monitor them: Yes, "monitor event logs" is a recommendation we gave last year and have already recapped above. However, we find that many organizations focus these logging efforts on network, operating system, IDS, and firewall logs and neglect remote access services, web applications, databases, and other critical applications. These can be a rich data set for detecting, minimizing, and investigating breaches.

Define "suspicious" and "anomalous" (then look for whatever "it" is): This is admittedly vague, but—in truth—generalizing what this entails in order to prescribe something for everyone would counteract the point. During this report, we have discussed extensively the increasingly targeted and sophisticated nature of a small but damaging subset of attacks. These typically occur within organizations that process or store large quantities of data valued by the criminal community; they are the quintessential Targets of Choice. Organizations falling under this classification should prepare to defend against and—especially—detect very determined, well-funded, skilled, and targeted attacks. Consider the "big breach" scenario highlighted in this report: The attacker gains access to the corporate network (potential warning sign number one), explores around a while (sign #2) to find and breach (sign #3) sensitive systems, installs malware (sign #4) that alters the state and processing of the system (sign #5). The malware stores data locally—often on a system that should not be storing any data (sign #6)—and the criminal returns periodically (sign #7) to collect the data and often does so via unusual ports (sign #8) and from known malicious IP addresses (sign #9).. Sound suspicious and anomalous to you? We think so too. Discover what is critical, identify what constitutes normal behavior, and then set focused mechanisms in place to look for and alert upon deviations from normality.

285 million records breached in a single year is a rather loud wake up call to an industry dedicated to protecting information. We can't afford to hit snooze and sleep in. Our task is not getting any easier; the sum total of information in the world grows continually and permeates everything we do and everywhere we go. While the majority of attacks remain rather mundane (let's take care of those first, by the way), the bad guys are adapting to our current protection strategies and inventing new ways to attain the data they value.

The good news is that we have more and more information at our disposal. A growing body of research aims to leverage real data and real results to help us better perceive reality and ensure that we are focusing on the right things. Sure, there's still a lot of noise out there but signal appears to be getting stronger. We hope our contribution is considered to be more signal than noise.

About the Verizon Business Investigative Response Team

Security breaches and the compromise of sensitive information are a very real concern for organizations worldwide. When such incidents are discovered, response is critical. The damage must be contained quickly, customer data protected, the root cause found, and an accurate record of events produced for authorities. Furthermore, the investigation process must collect this evidence without adversely affecting the integrity of the information assets involved in the crime.

The IR team has a wealth of experience and expertise, handling roughly 600 security breach and data compromise cases in the last five years. This includes a substantial proportion of all publicly disclosed breaches as well as many that have never been released. This caseload represents a large proportion of total known compromised records during this timeframe and many of the largest breaches ever reported.

During such investigations, the team regularly interacts with governmental agencies and law enforcement personnel from around the world to transition case evidence and set the stage for prosecution. In addition to security breach and data compromise cases, the IR team provides services such as litigation support, eDiscovery, expert witness testimony, chain-of-custody, mock-incident training, and incident response program development.

The expansive statistical data set generated through these activities offers an interesting glimpse into the trends surrounding computer crime and data compromise.

yr puvsser vaqrpuvssenoyr

www.verizonbusiness.com