Conflicts Among the Pillars of Information Assurance

Kelce S. Wilson, *BlackBerry*

Interactions between the five pillars of information assurance can be problematic. Measures taken to further the goal of one pillar are often blind to the needs of another pillar. The author explores such interactions using graphical forms to better represent conflicts.

The five pillars of information assurance—availability, integrity, authentication, confidentiality, and nonrepudiation[1]—aren't independent. Certain combinations can form pairs of differently focused requirements vectors, such that efforts aligned with advancing a goal of one pillar can frustrate or interfere with the goal of a different pillar. Availability, in particular, introduces conflicts with at least three of the other four pillars—confidentiality, integrity, and authentication. Meanwhile, confidentiality and integrity are different facets of a shared concept—the concept of controlled information access. Here, I explore these pillars and their interactions using graphical forms, which represent the pillars as orthogonal vectors in various 2-dimensional regions or as axes in regions representing each requirement.

Availability vs. Confidentiality

The aim of availability, at the highest level of abstraction, is to ensure timely and reliable access to data. When described in the contexts of confidentiality and authentication, access refers to reading the data (retrieval). However, when described in the context of integrity, access refers to writing the data (modification). Therefore, the pillar of availability can have different meanings at lower levels of abstraction, based on the context.

Maintaining adequate access to data is typically more of a challenge when storage-system or channel reliability is low. So the minimum required effort to ensure availability often varies according to the expected storage-system or channel reliability levels. Greater effort to maintain availability is needed when the reliability is

 Published by the IEEE Computer Society

low. The four-quadrant chart in Figure 1 illustrates this, using the vertical axis to represent "storage-system or channel reliability." The goal of availability, which is to preserve access, is indicated near the top of the chart, where reliability is marked as low.

Confidentiality, meanwhile, depends on the level of user authorization that the information owner allows. To illustrate this dependence, Figure 1 shows a horizontal axis, labeled "user authorization," which varies from high to low in the left to right direction. The goal of confidentiality is to prevent users who lack proper authorization from accessing data, which might seem to directly oppose the goal of availability. However, the depiction in Figure 2 illustrates the different goals as orthogonal, rather than directly opposing.

This notional orthogonality is a result of an interesting subtlety in the definitions of the different pillars' goals. For availability, the term "preserve" means to counter failings caused by storage-system or channel reliability—not failings in user authorization. For confidentiality, however, the term "prevent" means to stop disclosures when user authorization fails, rather than when a storage system or channel fails. Despite the initial appearance of clarity in the language, the actual interaction between availability and confidentiality is considerably more complex.

Figure 3 illustrates the different goals on the same chart. Although the term "access" might be consistently defined, the terms "preserve access" and "prevent access" have, as just noted, fundamentally different contexts. With this in mind, it's easier to interpret Figure 3 as illustrating four different situations:

- Both system reliability and user authorization are high, which reduces the need for expending significant efforts to preserve access against system failures and prevent disclosure.
- System reliability is low, while user authorization remains high, triggering the need to take measures to preserve access for authorized users.
- System reliability is high, while user authorization is low, triggering the need to take measures to prevent access by unauthorized users.
- Both system reliability and user authorization are low. Although no efforts are needed to preserve access for unauthorized users, efforts are needed to prevent access.

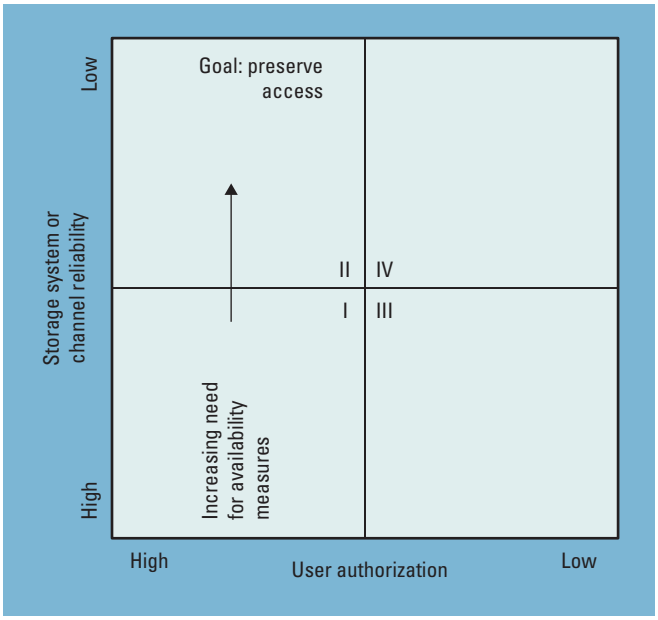

Figure 1. The "availability" goal in the context of confidentiality. The goal, which is to preserve access, is indicated near the top of the chart, where reliability is low.

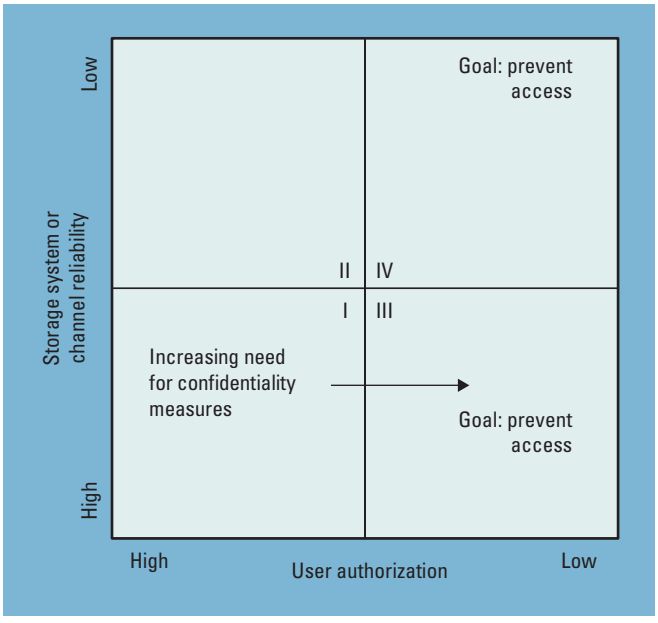

Figure 2. The "confidentiality" goal, which is to prevent unauthorized users from accessing data. The availability and confidentiality goals are orthogonal rather than directly opposing.

Unfortunately, because some measures taken to further the goal of one pillar are blind to the needs of another, efforts that mitigate reliability failure costs might be agnostic to user authorization,

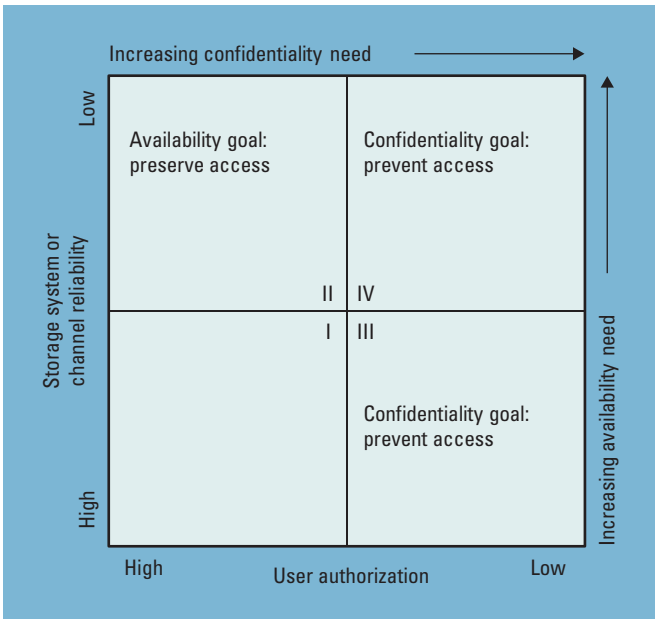

Figure 3. The orthogonality of availability and confidentiality. Although the term "access" might be consistently defined, the need to "preserve" and "prevent" access present fundamentally different contexts.

potentially frustrating user authorization enforcement efforts.

For example, some data backup schemes typically increase the number of data copies as risk of loss increases for unreliable systems, rather than decreasing the number of copies, to reduce theft opportunities by unauthorized users. In this manner, unintended consequences of availability measures can impede confidentiality measures. The data backup copy, created to ensure that data is still available in the event of the primary storage system's failure, presents a theft opportunity, which can compromise the data's confidentiality even if unauthorized users can't access the primary data storage system.

How does this threat materialize? Consider a military system for the US Department of Defense (DoD). DoD Instruction No. 8500.2 contains the requirement, "CODB-3 Data Backup Procedures," which appears in a "continuity" subcategory of the availability requirements. The requirement states that "data backup is accomplished by maintaining a redundant secondary system, *not collocated*, that can be activated without loss of data or disruption to the operation" (emphasis added).

There's an additional requirement (COBR-1) for "appropriate physical and technical protection of the backup." However pure the intent or stellar the competence of the system administrators and security operators, any noncollocated facility will become a potential alternate target for theft or tampering. Additionally, the transit channel from the primary site to the secondary site could become compromised. This presents two additional vulnerabilities that could be eliminated without a requirement for maintaining off-site copies of data. If a malicious party can't break into the primary site, budget constraints or personnel shortages might render an intended "appropriate" level of protection into at least two potentially easier theft opportunities.

The reverse condition is also true: unintended consequences of confidentiality measures can negatively affect availability. For example, an account lockout, resulting from a mistyped or incorrectly remembered password, can deny a user access to information in a timely manner. The delay introduced by waiting for a system administrator, who might be on a lunch break or attending to other duties, to reset the account and permit access, degrades the timeliness aspect of availability—at least during the period of the account lockout. It's important to note that, with this paradigm, availability and confidentiality aren't directly conflicting; rather, there's a potential for conflict if measures aren't properly implemented.

Another conflict between confidentiality and the timeliness aspect of availability can be demonstrated with security issues introduced when using logic-controlled shutdown procedures in mobile computing devices, such as cell phones. The logic-controlled procedures deactivate selected functions in a specified order but preserve other functions—such as timers, schedulers, and possibly some transceiver operations—to accelerate resumption of use, improving timeliness. Unfortunately, if malicious logic alters an insecure system's operation,[2] a tampered shutdown procedure could render the system susceptible to remote control, potentially putting confidentiality at risk. Sensors on a compromised device—such as audio, optical, location, and radio sensors—could be surreptitiously activated unless a hardware component, such as a mechanical switch, physically interrupts operation of the transceiver or sensor.

Another potential confidentiality measure, to prevent real-time access to a cellphone's sensors by eavesdroppers, is to break network registration

by placing the phone in a conductive, fully enclosed, RF-shielding holster that blocks signaling between the phone and serving base station. This also negatively affects the timeliness aspect of access by requiring reregistration with the network when the mobile device is removed from the holster. The reregistration introduces a delay that wouldn't be encountered in a holster that didn't block RF signaling.

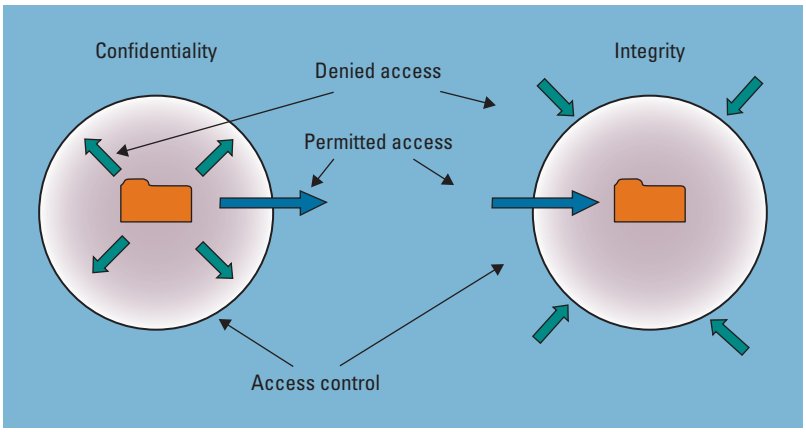

Figure 4. A comparison between confidentiality and integrity. The primary difference is whether the access control provides containment or shielding, respectively.

Confidentiality vs. Integrity

Figure 4 illustrates a notional comparison of confidentiality and integrity as complementary aspects of controlling access. For confidentiality, "access control" means to permit or deny access for retrieving information, whereas for integrity, it means to permit or deny access for altering or writing information. The decision to permit or deny access for both confidentiality and integrity depends on authorization.

Confidentiality measures function properly if authorized users can retrieve information, while attempts by unauthorized users are denied. Similarly, integrity measures function properly if authorized changes can be made to the information, while unauthorized changes are prevented or reversed. The primary difference, according to the notional illustration of Figure 4, is whether an access control provides containment or shielding.

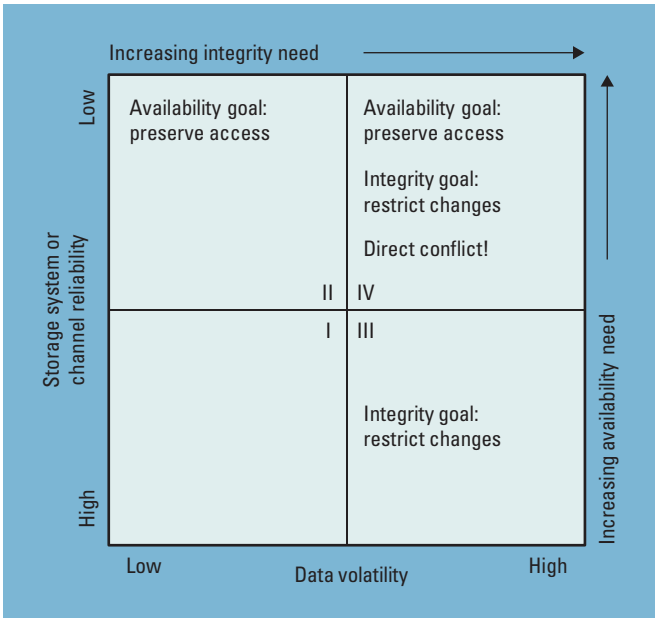

Figure 5. The orthogonality of availability and integrity. The horizontal axis no longer explicitly addresses only user authorization. It now references data volatility as well, and the highest-risk quadrant (IV) includes an availability goal of preserving access.

Availability vs. Integrity

The relationship between the pillars of availability and integrity is similar to that of availability and confidentiality. Both relationships hinge on interaction between authorization, as allowed by the information owner, and some form of access.

The goal of availability is to ensure timely and reliable access to data, but in the context of integrity, "access" means the ability to modify, rather than retrieve, data. Figure 5 illustrates a four-quadrant chart in which availability can be paired with integrity (somewhat equivalent to Figure 3) to illustrate a notional orthogonality of the different pillars.

In Figure 5, the horizontal axis has a nominally reversed direction of low to high, from left to right. This might appear, at first glance, as opposite the horizontal axis in the earlier series of Figures 1–3, but it's actually consistent with respect to the figure quadrants. Quadrant I is still the lowest-risk quadrant, because low data volatility provides friendly environments for integrity measures, similar to how high user authorization provides friendly environments for integrity measures.

The horizontal axis no longer explicitly addresses only user authorization, because integrity measures can counter additional threats,

including alteration of information content from entirely within a system. The horizontal axis now references data volatility, which includes the likelihood of alteration by a user, among other threats. Additionally, quadrant IV, which is the highest-risk quadrant, now includes an availability goal of preserving access. This is because user authorization isn't a specific issue in Figure 5. So, although this figure is similar to the earlier series, it does have some notable differences.

The high-risk quadrant, in which storage system or channel reliability is low and data volatility is high, indicates a direct conflict. One scenario in which availability measures might frustrate integrity is if a tampered backup copy that's used to compensate for a reliability failure is used as a false baseline to restore data to an incorrect state. So an attacker, who might not be able to directly alter information in the primary system, could attempt to access a backup copy that might be easier to access. He or she could then alter the backup copy and find a relatively easy way to sabotage the primary system. The system administrators, ostensibly working to assist the information owner, are the ones who actually—although unwittingly—introduce the tampered information into the primary system that's now improperly trusted by the users.

A reverse scenario is also easy to conceive, in which integrity measures can frustrate availability. False alarms happen, and the more sensitive a security system is intended to be, the more often a false alarm is likely to occur. With a low-reliability storage system or channel, data errors are more common. A genuine error could be mistaken for a tampering event, triggering a false alarm. In such a situation, data might be locked down to prevent further propagation of changes. If a properly authorized user had been attempting to make some allowable changes to the data, the lock down would impede the user's ability to distribute those changes until the security alarm had been investigated and resolved. Both availability measures can negatively affect integrity, and integrity measures can negatively affect availability, which is similar to the finding for the pairing of availability and confidentiality.

Availability vs. Authentication

Authentication measures are designed to establish the validity of a transmission, message, or originator, or to help verify an individual's authorization to receive specific categories of information. Proper authentication provides confidence (for the recipient) in a message's validity or its purported author. Of all the pillars described in relation to availability, authentication is likely to be the most directly contradictory.

Authentication can become a bottleneck, choking availability and potentially causing availability failures, while authentication mechanisms are in-process and yet to complete. This is because information that's awaiting authentication shouldn't be used for certain purposes and is therefore of little more value than information that hasn't yet been received.

A readily-understood example is to contemplate the situation of an automated teller machine (ATM) user, who has properly swiped the bank card, entered the correct PIN, and indicated the amount of a desired cash withdrawal. If the communication channel between the ATM and bank is interrupted prior to the ATM's receiving an authentication message from the bank confirming the cash withdrawal transaction, the ATM controller won't use the information from the user—or even any prior information received from the bank—to dispense cash to the user. Apart from perhaps some incident-logging, the ATM will operate as if the user had never entered the PIN or withdrawal amount. The availability of the information that the ATM had received is rendered effectively useless.

Availability measures can present additional opportunities for forgery and false representations, thereby increasing the risk of a false authentication. This might be inherent, because alternative or parallel sources or channels present additional vulnerable points of attack. Consider, for example, a data vendor with global distribution obligations that prepositions (or copies) data to globally dispersed sites near important markets to speed up delivery times despite infrastructure limitations, such as slow speeds and frequent outages. Each prepositioning site must provide an authentication scheme to enable recipients to trust a copy of data originating from that site. The duplication of authentication schemes, whether implemented by duplicated credentials or additional credentials requiring trust, presents opportunities for theft or spoofing in excess of a system having only a single site from which originating data requires trust.

Understanding these conflicts can help in analyzing existing information assurance measures to identify situations in which mitigation of one type of risk might increase another. Additionally, proposed measures should be analyzed to determine whether risk trade-offs are acceptable. Here, I highlighted four of 10 possible interactions. In the future, I hope to investigate the other six interactions among the pillars.

References

1. "National Information Assurance (AI) Glossary," Committee on National Security Systems (CNSS) Instruction No. 4009, 26 Apr. 2010; www.cnss.gov/Assets/pdf/cnssi_4009.pdf.
2. "DoD Directive No. 8500.2," US Department of Defense, 6 Feb. 2003; www.dtic.mil/whs/directives/corres/pdf/850002p.pdf.
3. Y. Kane, "How I Spent My Summer: Hacking into iPhones with Friends," *Wall Street J.*, 7 July 2009; http://online.wsj.com/article/SB124692204445002607.html.
4. "DoD Directive No. 8500.01E," US Department of Defense, 23 Apr. 2007; www.dtic.mil/whs/directives/corres/pdf/850001p.pdf.

Kelce Wilson *is a patent attorney for BlackBerry in the telecom industry and was previously an engineer for the Department of Defense, working in software security, surveillance radar, automatic target recognition, stealth aircraft technology, and satellite control. Wilson received his PhD in computational electromagnetics from the Air Force Institute of Technology. Contact him at kwilson@softwareipattorney.com.*