Thoughts and Questions

on

Common Criteria Evaluations

Kenneth G. Olthoff

National Security Agency

ABSTRACT:

The current United States scheme for evaluating products using Protection Profiles, Security Targets, and the Common Criteria takes an “all or nothing” view of evaluations, which may be also present in other nation’s schemes. While the author acknowledges the need for a binary grade on whether the profile has been met, he questions whether we should take a hard line stance on the use of the evaluation results in only pass or fail terms. In this paper, The author lays out why this may be an issue to be concerned about, and encourages analysis of the scheme for other unanticipated consequences.

One of the central points of the current strategy based on the international Common Criteria is the thesis that the Common Criteria will advance the state of security in two ways. It will encourage various parties to write Protection Profiles outlining their needs and desires, and it will push vendors to meet the resulting Protection Profiles. The theory proposes that as users profile desired capabilities that are not currently available, the vendors will attempt to gain market share by rising to the challenge. It is presumed that vendors will strive to be the first one to meet the profile in the hope of making their products a de facto standard. This chain of reasoning may be faulty, given the currently stated interpretation of the Common Criteria system in the U.S., which limits the use of information on failed evaluations.

Let's go through component parts of the Common Criteria, and the various ways we envision they will be used, at least as currently espoused by the proponents. There are two basic vehicles for describing the characteristics of a product or a system using the vocabulary of the Common Criteria, and then evaluating an instantiation of those characteristics. One is a Protection Profile, which is an implementation independent description of requirements. The other is a Security Target, which is an implementation specific description of the functionality of a particular product or system.

A Protection Profile describes a set of requirements that are specified with the aim of countering specified threats in a specified environment. The Protection Profile may not describe the optimal solution, but it is anticipated that it will be consistent, correct, and complete. In other words, it will not be self-contradicting. It will contain all the pertinent information to adequately talk about the problem space it seeks to address. It will not contain blatant errors of logic in relating the various pieces of the puzzle together - it will not state a requirement for a socket wrench as a means of hammering nails.

It is anticipated that a Protection Profile may be written by any of several parties. A Protection Profile may be written by a user community as a means of stating a need that is not adequately met by the current offerings on the market. A Protection Profile might also be authored by an accrediting body such as a government, industry group, or insurance firm. This might be done as a means of standardizing for interoperability (you must be able to interface according to these standards, or you won't be able to fit into our architecture). It also might be done to set a minimum standard for protection (you must have this level of protection, or you can't join our network, or get an insurance policy, or...). It is likely that both aspects will come into play to some degree.

What is not as commonly realized is that it may be written by a vendor, in hopes of putting an "open source, non-proprietary" spin on "implementation independent" requirements. The clever vendor might first describe their product in Protection Profile format, perhaps with the help of key customers. He would then write the product-specific Security Target in a way that points back to the Protection Profile. Not surprisingly, the product matches the requirements perfectly. The vendor might do so hoping that a customer group would adopt the profile the vendor has written, rather than go to the trouble and expense of writing their own. Such an arrangement might even be worked out in advance between a vendor and some preferred customers, with the customers using their influence with their peers to get the vendor's profile adopted as the

standard for some portion of the market. There has already been one instance of this strategy observed, though the results have not yet completely played out.

A Security Target by itself, being inherently product specific, would not have this patina of impartiality, and would not be as useful to the vendor in appearing to advocate open source, while still locking a market segment into what is an effectively proprietary solution. Yes, the vendor's competitors may still build to such a Protection Profile, but it will likely be written to limit their flexibility as well as their opportunity for product differentiation, giving the original vendor a substantial head start. This strategy may entail more work for the vendor, as well as some schmoozing of key clientele, but the benefits of such a strategy could be significant.

This is not to say that such a built-in advantage for a clever and resourceful vendor should not be allowed. It is a possibility we must be aware of, and one which we should factor into our projections of how the Common Criteria scheme might play out, and whether it will have the outcomes we desire in the long term.

The Security Target, by contrast, is implementation specific, and is the document which product evaluations are conducted against. Thus, the Security Target format will not be used to state requirements. Clearly, if the user or accreditor communities already have a solution in hand, they have little need to retroactively document and evaluate it using a Security Target. The vendor, however, does have a marketing use for such a document. It is a means by which the vendor can describe his specific product in the commonly understood language of the Common Criteria. The vendor can also have his product evaluated to provide potential customers with the independent testimony of the testing lab as to the truth of the claims he makes about his product.

There will almost certainly be vendors who will write a Security Target, and tell the user community "Here's what we've got, let us know if you're interested" rather than chasing the latest and greatest Protection Profile. That said, the marketing dynamics pushing a vendor to go the Protection Profile route, rather than using the Security Target alone, should be taken into account as a possible factor in our projections and strategies.

Given the outlines of the possible usage of Protection Profiles and Security Targets, how might these documents be used to advance the state of the practice or the state of the art? Since this has been one of the anticipated benefits of the Common Criteria scheme which has been most often cited, let us spend some time examining how it might play out in the future, based on the incentives inherent in the system.

Given that the Protection Profile is the anticipatory document, while the Security Target is an expression of what has already been implemented, we can safely say that Security Targets will be of only minimal influence in driving the future course of the security marketplace. Our thoughts about the means by which the security market will be driven should instead focus on the interplay between the Protection Profile and the results of evaluations. Keep in mind that this emphasis on the Protection Profile is only for this discussion and that for other discussions the Security Target may be a more prominent factor.

Let us now consider the possible usage of the Protection Profile using a hypothetical situation for

illustration purposes. Note that this example ignores the other sources of information and research about products that might be available in the "real world." The point of this discussion is to see how the Common Criteria scheme works. If the benefit of the Common Criteria is to provide a formal review by independent third parties, it should provide those benefits directly. Information gleaned outside the system may be presumed to be suspect. If it such outside information were sufficient, there would have been no need for the Common Criteria in the first place. Granted, this reduces a complicated matter to simple black and white, but it is useful for highlighting the question at hand. Further detailed analysis is left to the reader, and the author freely admits that this example is structured to make a point, rather than being a perfect reflection of all the factors involved.

Imagine we have some type of black box that the community cares about. If you wish, think of it as a firewall, a guard, an operating system, or whatever. The details of what the box is or does really don't matter. Let us further say that this type of black box is a type of device that has been around for a good length of time. It is not a new concept, but neither is it a product category where no further innovation is possible. There are still new developments that pop up from time to time, some of which advance either the state of the art or the state of the practice. Let us also assume that the state of the practice (the techniques commonly in use or available in products) considerably lags the state of the art.

It is currently contended that one of the major beneficial functions of the whole Common Criteria plan is that those who write Protection Profiles will be able to drive the market. Let's take the hypothetical situation laid out above and examine the incentives presented to the various participants by the structure of the current system and by the marketplace.

The user and accreditation communities are variants of the same set of interests. These parties may have a desire to push the vendors to provide more functionality, and they may choose to use the Protection Profile to do so. That said, which strategy should they pick for the optimal result?

The user community can write profiles they know can be met by currently available products. This will set a minimum standard. By doing so, however, they get no improvement above the current state of the practice, and depending on how many products meet the levels they establish, they may not even get any substantial product differentiation. To use an automotive analogy, specifying seatbelts as a minimum safety standard neither advances the current state of auto safety practice, nor does it provide any differentiation - all cars produced or sold new in the U.S. since the mid 1960s have had seatbelts installed. Such a profile feature would document seatbelts as a minimum requirement, but doing so would add no incentive to improve.

In the case of a Protection Profile which calls out standards that all competing products easily meet, the user community will at minimum get the benefit of independent evaluation of the products against their profile. They may not, however, get much new information for the expense and effort expended. This will be especially true if the Protection Profile authors research available products to ensure that the standard is not set too high. It is entirely possible that in order to get sufficiently precise information upon which to base their Protection Profile, the user community will have already gathered nearly as much information as the formal evaluation will provide. In such cases, the evaluation process boils down to buying a seal of approval, or perhaps document-

ing known facts to demonstrate “due diligence” or to avoid liability. It does not provide any impetus to advance the state of the art or the state of the practice.

The more interesting situation comes when the Protection Profile writer wishes to push beyond either the state of the practice or the state of the art. In these cases, the user must weigh the cost against the potential benefit. Clearly, if the user writes a Protection Profile with which to push the vendor to greater efforts, the trick will be to push, but not push too far. One wants to write a Protection Profile that will inspire the vendors to produce products with new or better functionality. If one sets the standard too high, though, the vendor may either not be able to reach it, or may choose to not try, deeming the cost too high for the perceived benefit. In either case, given the current scheme, the Protection Profile writer will have wasted his efforts. He will either have evaluation reports which tell of failure, with little additional useful information, or will have no evaluations whatsoever to look at because the vendors balked at trying to meet the overly ambitious Protection Profile.

Assume the following functions, all of which our hypothetical profile writer finds desirable.

FUNCTION	STATUS	DEPENDENCIES
A	Widely available, but of use only to a small portion of the market	None
B	Widely available	C
C	Widely available	B
D	Theoretical -D may enhance the effectiveness of B and C, but the actual usefulness has not yet been proven either by testing or by abstract reasoning.	B,C
E	Known, but not widespread	none
F	prototyped	E
G	Known idea, implemented in some products	none
H	Limited application	none

Given this hypothetical set of functions, what combination might a Protection Profile writer pick? Functions A, B and C are pretty easy choices, as they are already the state of the practice. But what of the others? Should one go for E, which is state of the art, but not yet state of the practice? Does one take a chance on the theoretical D before F, G, or H, because of the

enhancement it provides to B and C? All of this is of interest, and the calculus of risk and reward needs to be developed. Remember, the writer wishes to push the market and increase the security of commonly available products, but not to set a standard that the market can't or won't meet.

A major factor in all of this is the aftermath of the evaluation. As the plan is currently laid out, there is a strong emphasis on the "all or nothing" nature of an evaluation. A product either meets the profile or it doesn't. Does this encourage the sort of market driving we claim to want? In our example above, consider the outcome if our hypothetical profile writer creates a Protection Profile which calls for functions A, B, C, D, E, and G. Let us look at some possible scenarios for vendors' responses.

Vendor 1 - Has a product that performs Functions A-C and F, and has a large percentage of market share. She is the market leader. Her Company has even generated a Security Target for her product, and has had the product successfully evaluated. She announces plans to submit for evaluation against the new profile at some point in the future, implying that the product should pass the evaluation, despite knowing the product does not support functions E or G.

Vendor 2 - Has a product that can meet all but Function A, which he has already made a business decision to not include in the product, due to the limited interest. Does not submit the product, knowing the product will fail on Function A.

Vendor 3 - Her product fails on item D. The product is otherwise very strong.

Vendor 4 - His product is totally without technical merit, and fails miserably. Despite this, Vendor 4 notes in his advertising that the product was submitted for evaluation against this rigorous profile and notes to the user community that nobody else has been successfully evaluated against the profile.

There are many more possible scenarios, but these four present a few key points.

Vendor 1 shows us that while one may limit the usage of evaluations in advertising, it is hard to fight against speculative statements about what the outcome might be if a given product were to be evaluated. It is clear that an established vendor will make use of the evaluation system when it fits the vendor's aims, and may gain some benefit without playing at all in other cases. Since the Common Criteria is an international standard, it will not be possible to restrain discussion of it in the same way that one might with a proprietary standard. For example, one cannot claim to meet Underwriters Laboratory ("UL approved") standards, because UL owns the right to the use of their service marks, and the standards themselves are not widely known as a separate entity.

Vendor 2 shows us the limit placed on innovation by an "all or nothing" dictum. Consider the possible result if instead of a "pass/fail" grade, Vendor 2 were able to submit a product for evaluation against a profile, and be allowed to make honest use of the detailed results, even if the overall grade was a failing one. In such a case, Vendor 2 might be more likely to submit the product for evaluation. In an all or nothing scenario, the lack of Function A makes the whole discussion moot, no matter how well the product performs on the rest of the evaluation. If Function A was a "nice to have" item rather than a true necessity, Vendor 2's product might be

of interest. Vendor 2 has no way of knowing that, though, and the user has no way of knowing that Function A was the only thing Vendor 2's product lacked, because it never gets submitted.

Vendor 3 and Vendor 4 point up a similar concept. In cases where nobody meets the profile, limiting the "official" information to a pass/fail grade prevents the user from differentiating between a strong product which only failed on one point of lesser importance, and a horrible product which failed on almost everything. The more that a given Protection Profile "pushes the envelope, the more likely it is that this problem will come into play.

Vendor 3 has no vehicle by which to provide independent documentation of the functions that are properly implemented. The official result report lists "fail" and goes no further. The data on how it failed is provided to the vendor, but as the scheme currently is envisioned, that information is not intended to be quoted to customers as endorsing the portions of the evaluations which were successfully passed.

Vendor 4 clearly means to imply parity with his competitors, and the lack of detailed data available to the customer base beyond the "pass/fail" evaluation allows him to continue to do so. The evaluators are not authorized or intended to serve as a "truth squad" on claims of this sort. The vendor is telling the literal truth - nobody has passed. There is no means to authoritatively compare how close the various products came to passing, and this works in the favor of the more unscrupulous marketers.

The user can potentially go back and write another, less rigorous profile, but without reliable feedback as to exactly what portions of the profile are causing the products to fail evaluation, the user has little to guide the revisions. Providing such feedback, though, would rapidly erode the stated "all or nothing" policy. It could be a long and costly process to rewrite the profile and pay for subsequent evaluations of multiple competing products until somebody can meet it. Functions that could be met may be deleted from the profile in the process of trying to get a profile that the market can meet. The lack of official feedback to the profile writers leaves them guessing as to what requirements to relax or delete. Even in a situation where multiple products meet a profile, disclosure of detail beyond pass or fail might be useful in determining the relative strengths of the competing products, even if only in areas where specific objective metrics are called out by the profile.

The situation is equally bad on the vendor side. Even though the vendor may be told in what ways the product is deficient, there is no means by which the vendor can present impartial documentation of evidence to the customer community of what is right with the product. In effect, while the vendor is told what specific tests the product passed, and which ones the product failed. The only part of the evaluation results that will be independently confirmed to a customer, though, is the fact that the product failed. The specific evaluation feedback provided to the vendor is not documentation that the scheme intends the user to rely on. That said, if detailed documentation is provided in order to allow the vendor to improve, one wonders by what legal mechanism the vendor is to be prevented from using or sharing that information as he chooses. It's also unclear how much or how little trust the customers will put in the vendor's description of the evaluator's feedback, given the policy of the scheme to neither confirm nor deny any details beyond the "pass or fail" result.

The current scheme implies that the evaluators will not be allowed to confirm to the profile writer or to other customers anything other than the fact that the product failed the evaluation. This emphasis on the "all or nothing" nature of the current evaluation strategy may be misguided in some cases. Clearly, if the profile reflects the stated needs of a user community, and at least one product passes evaluation, the details of the failures may be of limited interest.

If, however, there are no products which successfully meet the profile, it is in the best interests of both the user community and the vendors to allow dissemination and confirmation of the details of the evaluation results, if the vendor chooses to release them. Given a product which failed all tests and a product which failed only one test when evaluated referring to the same protection profile, the customer would definitely benefit from knowing which product had the better results, even if both failed. Even in cases where one product passed and one product failed by a small margin, the customer may wish to know this. A substantial price difference or the nature of the test that the one product failed may make the failed product the better buy for some applications.

While it is still possible for a vendor to withhold the details of an evaluation failure report, competitive pressure will tend to bring such details to the attention of the user community. The vendors whose products came very close will have incentive to disclose the full details, and to work with the user community toward a mutually acceptable solution. Also, vendors who see an unmet need may have incentive to submit to evaluation despite shortcomings in their product if the documentation of the evaluation can be used in such discussions with the customer despite the failure.

A hard line stance on non-disclosure of failure data and the pass/fail system may tend to drive vendors toward bypassing Protection Profiles altogether, or at best, only submitting products for evaluation against those profiles which present no possibility of failure. Rather than go through the expense of testing against multiple Protection Profiles, many vendors, especially the ones in market leading positions, may instead put forth their Security Target in a "Take it or leave it" stance. Alternatively, vendors may collude with key customers to have a Protection Profile written which steers the market toward their product. Those in a position of market dominance will be able to keep market share without playing the game at all.

CONCLUSION:

Depending on the outcomes one wishes to see, there are many different strategies we may implement using the Common Criteria framework. We in the security community must do careful analysis of the incentives and disincentives inherent in those strategies, and the actions to which the various parties may be driven to as a result. We must be especially careful to understand the economic and business aspects, as these are the prime drivers for the commercial vendors whose behavior we most wish to influence. The preceding analysis of the "pass/fail" approach, while admittedly contrived to present the issues with less subtlety than would occur in actual practice, is still representative of the sort of questions we should be asking. We should encourage more of this type of investigation of the incentives built into the system and the likely results, both intended and unintended, rather than blindly assuming things will work as we wish them to. By doing so we may then be able to more closely align the behavior drivers inherent in the system with our desired outcomes.