Trends in Government Endorsed Security Product Evaluations

Richard E. Smith
Secure Computing Corporation

rick_smith@securecomputing.com

Abstract

Government endorsed security product evaluations completed between 1984 and 1999 show a number of interesting trends. According to official lists maintained by the United States, United Kingdom, and Australia, 242 products were evaluated since the first U.S. evaluations in 1984. The evaluations show trends in the number of evaluations performed, which evaluations are performed and where, what evaluation levels are achieved, and what types of products are evaluated. The average number of evaluations completed each year tripled in 1990 and again in 1994 to reach an average rate of 30 per year. Evaluations based on the Information Technology Security Evaluation Criteria (ITSEC) have predominated since 1993 and accounted for 70% of the evaluations completed in 1999. Although the Common Criteria is the recognized international replacement for older national criteria, it only accounted for 17% of the evaluations completed in 1999.

Introduction

Government endorsed security product evaluations continue to take place even though the increasingly rapid pace of technological change in the Internet era would seem to leave them behind. A survey of official lists has uncovered reports of 242 product evaluations completed between 1984 and 1999. These lists were produced by the United States, United Kingdom, and Australia. In addition to citing evaluations performed in those countries, the lists included references to evaluations performed in Canada, France, and Germany. These included evaluations against the following:

- *Trusted Computing System Evaluation Criteria* (TCSEC) performed in the United States [1].
- *Information Security Technology Evaluation Criteria* (ITSEC) performed in Australia, France, Germany, and the United Kingdom [2].
- *Common Criteria* (CC) evaluations performed in the U.S., U.K., Australia, Canada, France, and Germany [3].

Security evaluations began in earnest when the U.S. National Computer Security Center (NCSC) initiated product evaluations against the TCSEC. The first edition of the TCSEC was published in 1983 and the final edition was published in 1985. The TCSEC defined a rigid set of functional requirements and assur-

ance requirements. The functional requirements identified the system's technical features and capabilities, like access control lists and classification labels. The assurance requirements identified specific activities that had to be performed and documents that had to be written to help provide confidence that the system worked correctly. At the low end, assurance requirements focused on testing and basic documentation, while at the high end they involved formal proofs of security properties. The TCSEC defined an ordered set of evaluation levels (C1, C2, B1, B2, B3, A1) in which the lowest level evaluation combined low assurance with limited functional requirements. Higher assurance systems had to provide richer sets of predefined security functions as well as higher assurance through more stringent test, review, and analysis requirements.

Regardless of a system's actual purpose and security requirements, the TCSEC required total compliance with functional requirements in order to earn an evaluation rating. As other countries developed their own evaluation criteria, many relaxed the rigid functional requirements while retaining the notion of achieving higher confidence through increasingly stringent assurance requirements. This led to the ITSEC in the late 1980s, which was adopted by several countries, primarily in Europe. The ITSEC defined seven different evaluation levels (E0 through E6) in terms of assurance requirements. To evaluate functional requirements, the ITSEC required the "sponsor" of the evaluation (usually the product's developer) to publish a *security target* document which identified the relevant security requirements. This allowed the sponsor to tailor the evaluation to address a product's actual capabilities instead of having to modify the product to include potentially expensive features that might not otherwise be required.

In the mid-1990s, U.S. efforts to improve or replace the TCSEC were combined with international efforts to improve the ITSEC. This yielded the Common Criteria, which was recognized as an international standard in 1999 (ISO/IEC Standard 15048:1999). Like the ITSEC, this new criteria defined a series of evaluation levels in terms of increasingly stringent assurance requirements (EAL 1 through EAL 7). The Common Criteria also incorporated the notion of *protection profile* documents, which capture specific sets of functional and assurance requirements to apply to specific types of products. For example, there are protection profiles for Internet firewalls, smart cards, and multilevel operating systems. In 1998, the governments of Canada, France, Germany, the U.K., and the U.S. signed a mutual recognition agreement so that many of the Common Criteria evaluations performed in one country will be accepted in the other countries. Recently, Australia and New Zealand also signed this agreement.

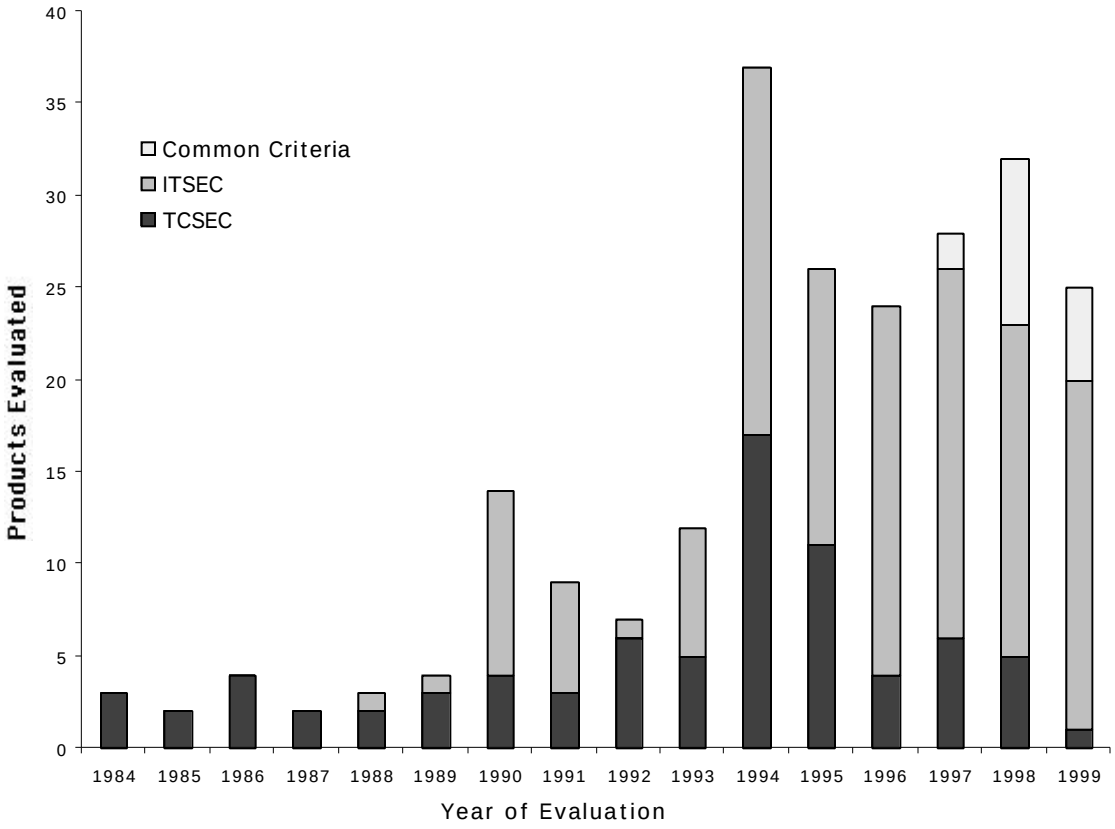

Figure 1: ITSEC evaluations dominate, despite the pioneering role of the TCSEC

Popularity of Different Criteria

Figure 1 illustrates the number and types of evaluations completed annually from 1984 through 1999. Note that there are two major jumps in the number of evaluations: 1990 and 1994. In both cases, the average number tripled following the jump. In the 1980s, an average of 3 evaluations were finished every year. From 1990 through 1993, the average was over 10 per year. Starting in 1994, the average was 30 per year.

The increases have primarily been in ITSEC evaluations. ITSEC dominated product evaluations almost every year during the 1990s. A typical year sees the completion of perhaps 5 TCSEC evaluations; this rate has been fairly constant except for a brief jump in 1994 and 1995. There are several reasons why TCSEC evaluations have been less popular than ITSEC with product vendors:

- Rigid functional requirements - it was noted earlier that TCSEC mandates a variety of technical features that a particular product might not otherwise need. ITSEC and Common Criteria evaluations can be tailored to match the product's existing functional capabilities, which simplifies the problem of making an existing commercial product comply with the evaluation requirements.

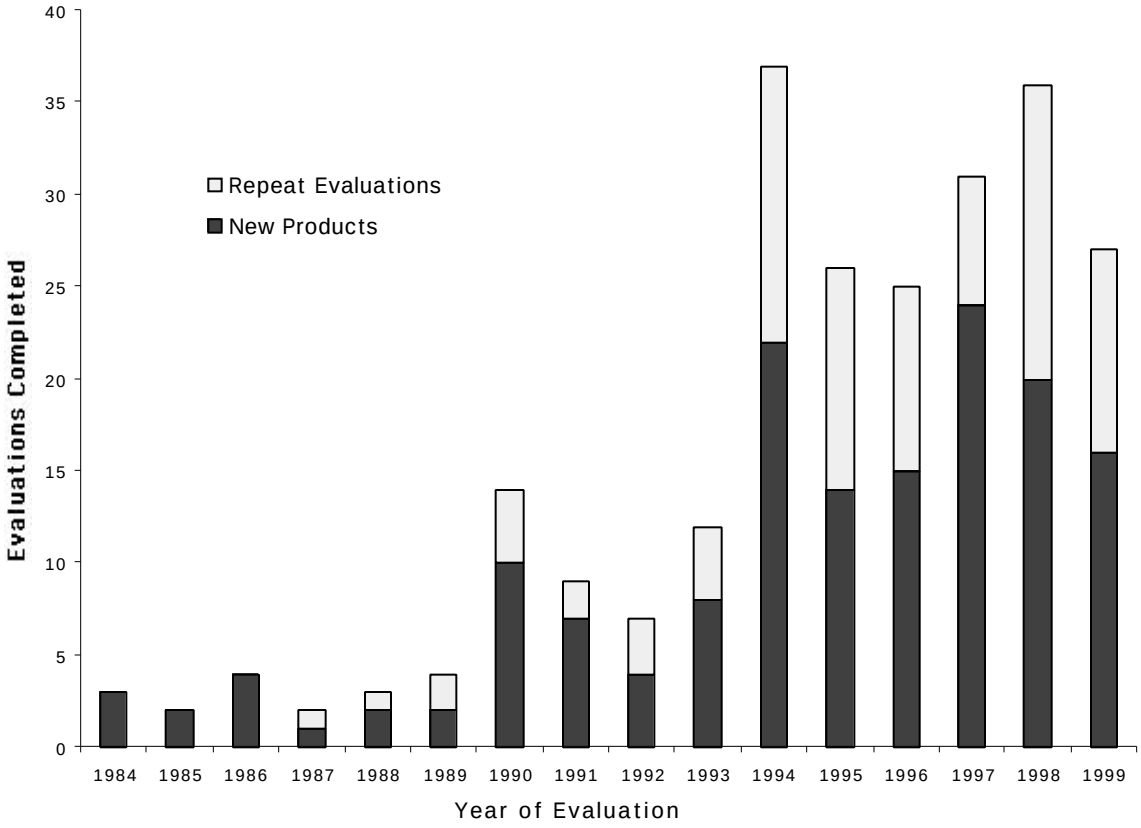

Figure 2: Over a third of all evaluations since 1987 have been re-evaluations

- Scheduling problems - Originally, all TCSEC evaluations were performed by the NCSC itself. Limitations of staff and other resources would explain the small yearly rate of TCSEC evaluations. The ITSEC pioneered the use of commercial laboratories to perform evaluations, a strategy the NCSC eventually adopted. The Common Criteria also relies on commercial labs.
- Cost - in the mid-1990s, a vendor informally estimated that a recently completed B1 evaluation had cost $2.5M. The commonly accepted cost estimate for a roughly comparable ITSEC E3 or Common Criteria EAL 4 evaluation is $1M.
- U.K. government mandate - Unlike the U.S. government, the U.K. government actually enforces its requirements to use evaluated security products. Before the Common Criteria was introduced, vendors had to participate in a U.K. ITSEC evaluation in order to sell products to the U. K. government.
- Grandfathering - Once a product has been evaluated under a particular criteria, it is usually less expensive to evaluate a revised and upgraded version of the same product under the same criteria.

Figure 2 compares the number of new evaluated products every year against the number of repeat evaluations. On average, 38% of the evaluations completed every year since 1987 have been repeats. There are two reasons for repeat evaluations. The first and earliest reason was the introduction of a new version of an earlier product. Between 1994 and 1996, Digital Equipment Corporation completed 3 different evaluations of its SEVMS product, each representing a different version. The second reason was that products occasionally had to comply with local procurement requirements that favored that nation's criteria. For example, Microsoft Windows NT, Version 3, received a TCSEC C2 rating in 1995 and an ITSEC E3 rating in 1996.

The Common Criteria, which is intended to replace ITSEC, accounted for 25% of the evaluations in 1998 and 19% of those in 1999. This relatively modest showing is probably because evaluation laboratories are new at planning, pricing, and executing Common Criteria evaluations. Vendors probably still use older criteria where possible, since it most likely requires less planning and other resources to redo a previous evaluation than to pursue one under the new criteria. Common Criteria evaluations should occur more often as vendors choose to pursue a single, internationally recognized Common Criteria evaluation instead of earning separate ITSEC evaluations in different countries. Furthermore, the NCSC has officially stopped performing TCSEC evaluations, so new product evaluations in the U.S. must now use the Common Criteria. Most countries that do ITSEC evaluations today have agreed to migrate to the Common Criteria, so it should eventually replace all other criteria.

Evaluation Levels Achieved

Figure 3 illustrates the evaluation levels achieved by evaluated products over the years. In order to compare these levels across the different criteria, we must identify which level in one criteria corresponds to which levels in the others. Figure 3 uses the mapping shown in Table 1. Keep in mind, however, that the TCSEC is significantly different from the ITSEC and Common Criteria. This mapping only reflects a rough correspondence between evaluation levels, and even experts disagree on what mappings make sense.

TCSEC	D	—	C1	C2	B1	B2	B3	A1
ITSEC	E0	—	E1	E2	E3	E4	E5	E6
Common Criteria	—	EAL 1	EAL 2	EAL 3	EAL 4	EAL 5	EAL 6	EAL 7

Table 1: The rough correspondence between evaluation levels of different criteria

A major goal of the pioneering U.S. evaluation program was to encourage computer systems vendors to build high assurance systems that achieved the highest evaluation levels. In the early days of evaluations, there was a belief (or

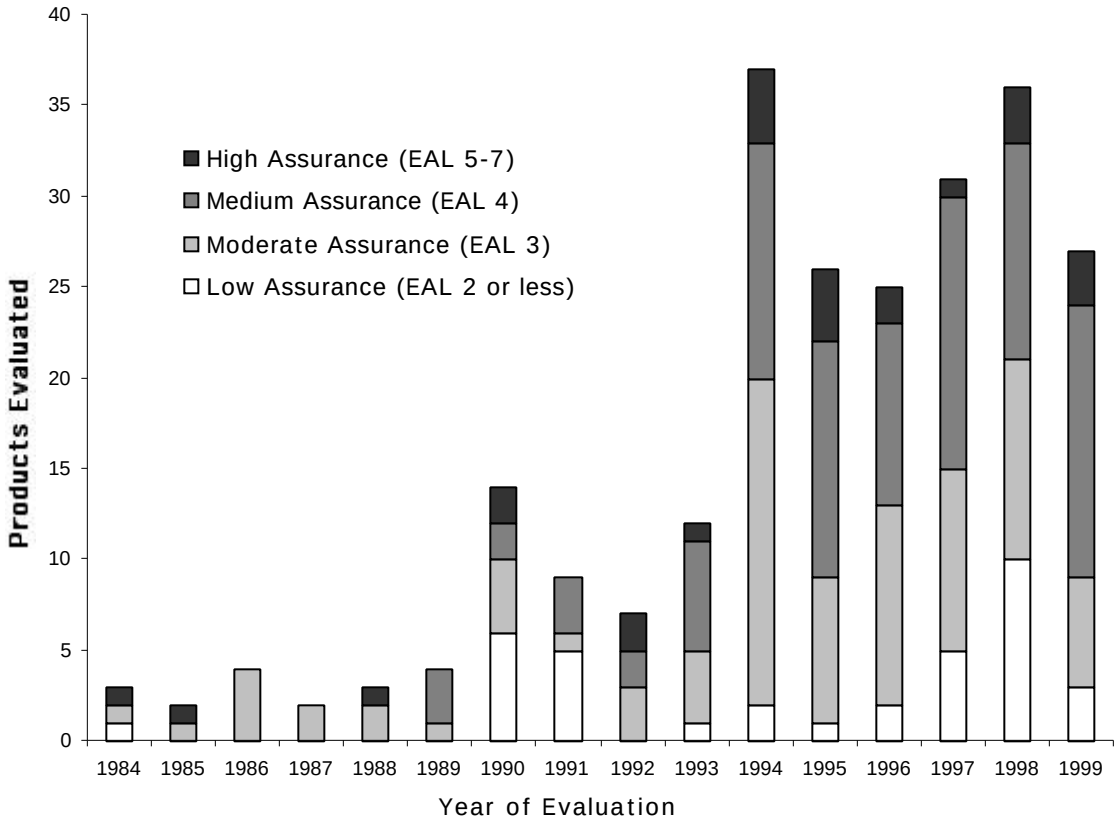

Figure 3: Most evaluations have achieved a mid-range (EAL 3-4) level of assurance

at least a hope) that major vendors of the time, like IBM and Digital Equipment Corporation, would upgrade their standard operating systems to meet TCSEC requirements and then do the extra work necessary for a high assurance evaluation. This never happened. Almost every system that completed a high assurance evaluation became a special purpose product for the military or intelligence community. The few exceptions were commercial failures, like Honeywell's MULTICS.

Low assurance products have principally been personal computer security packages and, more recently, Internet firewalls. Few products were evaluated under the TCSEC at low assurance levels, although a few "earned" a D rating through a failure to earn anything higher. Common Criteria EAL 2 mandates some basic documentation and requirements based testing, so it is a practical objective for a tolerably sophisticated software development organization. Although their development activities might already be more extensive than the minimums established by EAL 2, many vendors resist the costs of tailoring their development process to the peculiar needs of evaluators. It is worth noting that the U.K. government often requires a minimum evaluation level of E3 or EAL 4 when purchasing security products like firewalls.

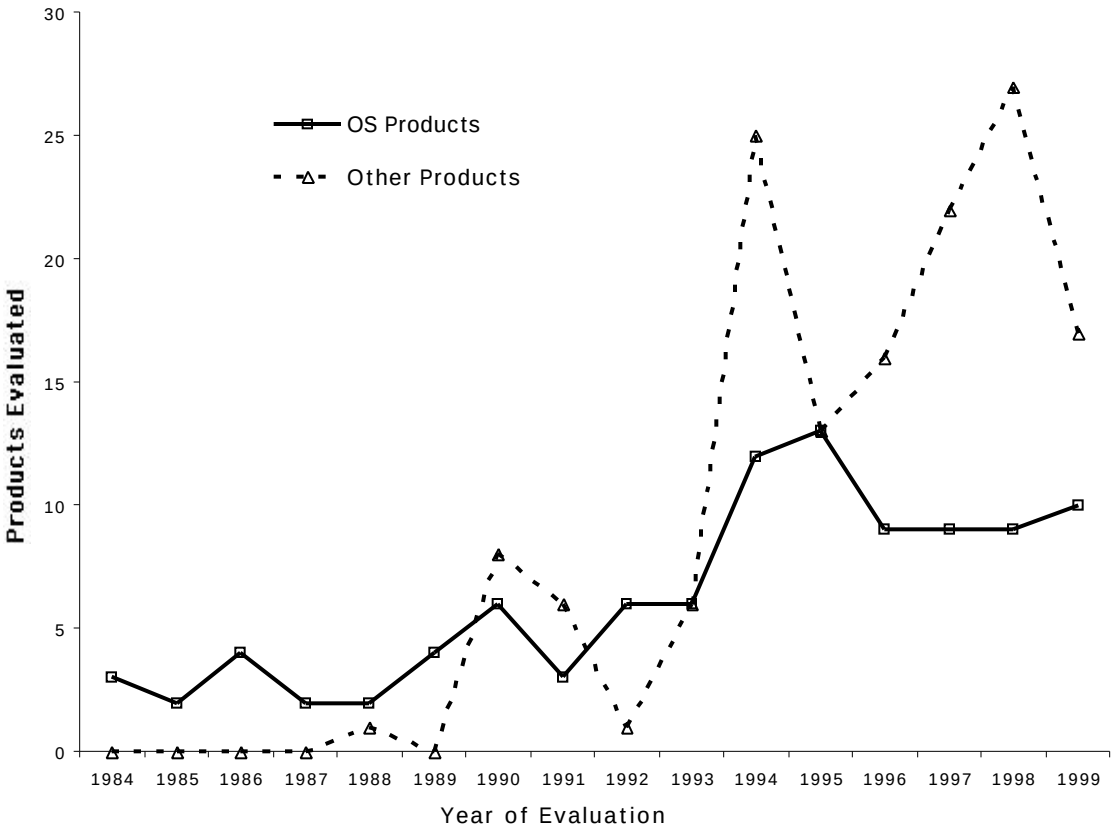

Figure 4: Operating system products no longer dominate evaluations

Evaluating Non-Operating System Products

The clear purpose of the original TCSEC was to establish security standards for operating systems and related access control products. As the community gained experience with security evaluations, it was obvious that OS evaluations were too narrow for many security situations. This led the NCSC to publish "interpretations" of the TCSEC to apply to networks and databases. This did not really address the problem as well as the ITSEC and the notion of security targets.

Figure 4 shows how non-OS products have come to dominate evaluations. In the 1990s the number of OS evaluations leveled off at a rough average of 8 per year. During the same period, other types of products accounted for an average of 14 evaluations per year. To some extent this undoubtedly reflects the popularity of firewalls, encryption devices, and other perimeter security devices. While a customer might not pay the necessary premium to populate a site with an evaluated OS, the same customer might pay a premium for evaluated equipment to protect the site's perimeter. In fact, the Common Criteria community

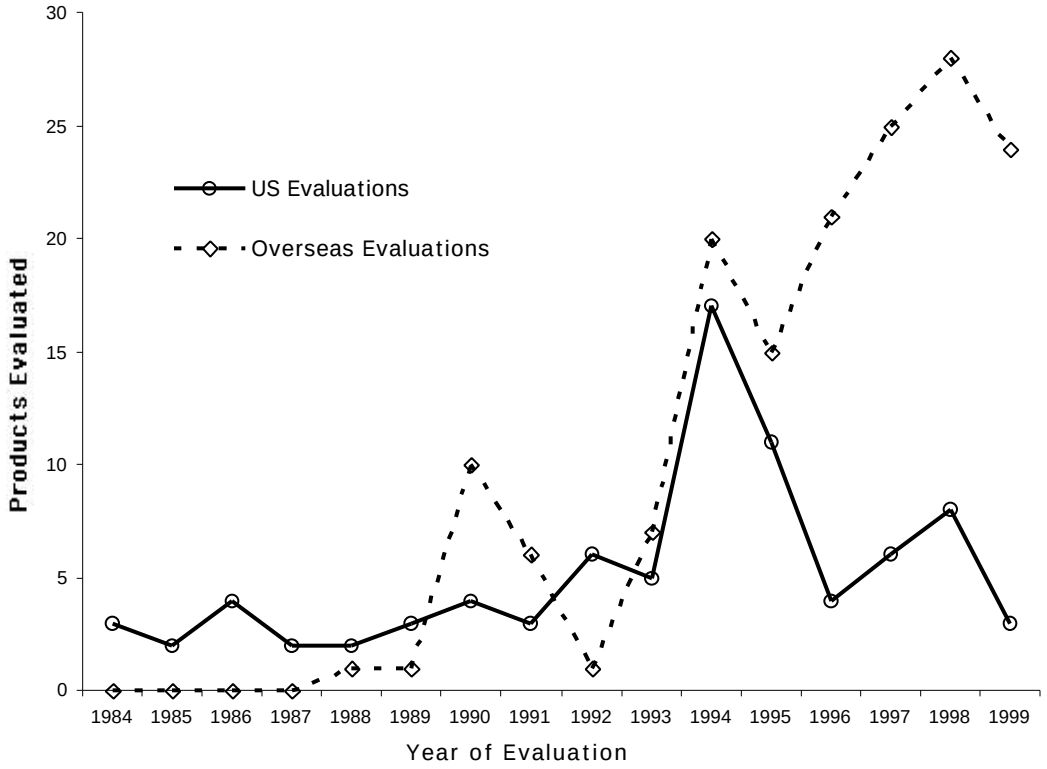

Figure 5: Despite early leadership, few evaluations are now performed in the U.S.

has published profiles for firewall products, and several firewalls have been evaluated under the ITSEC and the Common Criteria.

Evaluations Within the U.S.

A final trend worth noting, shown in Figure 5, is an interesting side-effect of the popularity of ITSEC. Despite the fact that the U.S. pioneered security evaluation and U.S. companies arguably dominate the international market in information technology, most product evaluations, including those of U.S. products, take place outside the U.S. Since 1993, less than 50% of all evaluations were performed in the U.S., and less than 25% of those since 1996. This is despite the adoption of the Common Criteria by the U.S. government, including the use of commercial laboratories to perform U.S. evaluations.

Although it may be too soon to judge the effect of the Common Criteria on U.S. evaluations, certain problems continue to discourage evaluations inside the U.S., even by U.S. vendors. First, there is the grandfathering effect: it is easier to re-evaluate a product using the same evaluation lab than to take the product to a different lab. Products that already carry ITSEC or Common Criteria ratings from an overseas lab are therefore more likely to pursue re-evaluation through the same lab. Cost, however, is a more significant problem. Initial es-

timates for Common Criteria evaluations by U.S. labs in late 1999 were between 100% and 200% higher than estimates from the U.K. It took a lot of negotiation and detailed costing to narrow the gap. Labs in both countries must follow the Common Evaluation Methodology, which should yield the same results regardless of the lab used. Therefore, costs should converge over the long term.

Summary

- Security product evaluations have grown with the adoption of the ITSEC and the use of commercial laboratories to perform the evaluations. Adoption of the Common Criteria in the U.S. provides similar benefits to vendors that desire a product evaluation within the U.S. itself.
- Re-evaluation accounts for a significant fraction of product evaluation activity. Since it's easier to re-evaluate a product through the same evaluation organization, patterns of evaluations within particular nations, criteria, and labs are likely to persist. However, the NCSC has phased out TCSEC evaluations, so products with TCSEC evaluations must now be evaluated under the Common Criteria in the U. S.
- Non-operating system products account for the larger share of evaluations as the marketplace has evolved a broad range of security products. The growth of firewall evaluations suggests an increased emphasis on evaluation of specialized security products instead of more general products.
- Vendors pursue the evaluation criteria and labs that offer the lowest costs. This is especially important for Common Criteria evaluations, since these evaluations are officially recognized by all participating nations.

Web Sites (as of June 2000)

Spreadsheet data:	http://www.visi.com/crypto/
Australia:	http://www.dsd.gov.au/infosec/pdfdocs/EPL.pdf
Common Criteria:	http://csrc.ncsl.nist.gov/cc/
UK ITSEC Evaluations:	http://www.itsec.gov.uk/products/
U.S. Evaluated Products:	http://www.radium.ncsc.mil/tpep/index.html

References

1. National Computer Security Center, "Trusted Computing Systems Evaluation Criteria," DOD 5200.28-STD (Ft. George G. Meade, MD: National Computer Security Center, December 1985).
2. "UK IT Security Evaluation and Certification Scheme: Description of the Scheme," UKSP01, Communications-Electronics Security Group, Gloucesterchire, March 1991.
3. Common Criteria Project Sponsoring Organisations, "Common Criteria for Information Security Evaluation," Version 2.1 (:Common Criteria Project Sponsoring Organisations, August 1999)