The U.S. GOSIP Testing Program

Jean-Philippe Favreau
Kevin L. Mills
J. Stephen Nightingale
National Institute of Standards and Technology

July 31, 1990

Abstract

At the 5^{th} International Conference on the Application of Standards for Open Systems Interconnection, Kevin Mills, representing the National Institute of Standards and Technology, described the challenge of establishing a test policy and procedures for the U.S. GOSIP. He said that to '. . . serve as effective tools for establishing GOSIP conformance and interoperability of [computer] systems, the policies, procedures, and techniques must be technically credible, must be accepted by vendors and buyers, must provide assurance of interoperability, and should serve as a basis for international recognition of national testing.' The challenge has been more difficult than we imagined, and yet, we have made significant strides to meet our goals. The following paper recounts our progress since March of 1989, describes the policy and procedures we now have in place, and looks ahead toward remaining challenges.

Introduction

In April 1989, sixteen months before the U.S. Government mandate to use Open Systems Interconnection (OSI) in procurement of new networks and major upgrades to existing networks, a large Federal Agency issued a procurement specification for a multi-year computer buy, including a requirement to conform to the *U.S. Government OSI Profile* (GOSIP) [1]. The agency sought our advice on GOSIP compliance, specifically: How can vendor claims of GOSIP compliance be substantiated? Our answers were far from convincing, and so, the agency, a voluntary user of GOSIP in advance of the OSI mandate, withdrew the requirement for GOSIP compliance.

Following this incident we conducted an analysis of the OSI testing situation and concluded that, unless the National Institute of Standards and Technology (NIST) acted, no credible means of substantiating GOSIP compliance would be available in time to support the U.S. Government OSI mandate beginning August 1990. Abstract Test Suites, where they existed, were fragmented and not publicly available. Although multiple suppliers of Means of Testing (MOTs) existed, no credible mechanism existed to assess MOTs against GOSIP requirements; no means existed for finding one MOT acceptable and another not. No program of evaluating and accrediting commercial GOSIP testing laboratories was planned. Numerous policy issues were unresolved, including requirements for first party versus third party testing and the role of interoperability testing; no forum was foreseen for shaping policy.

From April through November 1989, we defined a GOSIP Testing Program to permit Federal Agencies to substantiate claims of GOSIP compliance and to provide a forum for shaping GOSIP testing policy. On November 13, a proposed *GOSIP Conformance and Interoperation Testing and Registration* Federal Information Processing Standard (FIPS) [2] was issued for public comment, followed by a public meeting on November 27 to explain what we were doing and why. Since November, we have made progress on the details of GOSIP testing, including identifying Abstract Test Suites, developing methods to assess MOTs, establishing a program to accredit laboratories, defining the role of interoperability testing, creating a set of publicly accessible registers, and detailing a schedule synchronized with the U.S. Government OSI mandate. These topics are treated briefly below, followed by a statement of the U.S. GOSIP testing policy, before we pause for an assessment; and then we look ahead to some remaining challenges.

Abstract Test Suites

Our first priority was identification of Abstract Test Suites (ATSs) covering functionality in GOSIP, mainly based on the Stable Implementation Agreements from the OSI Implementors Workshop (OIW). An ATS is the basis for evaluating MOTs and vendor products. Ideally, an ATS should be produced in concert with a standard, follow the standard into implementation forums, and be maintained along with the standard. For most OSI standards, no standard ATS exists; two notable exceptions are X.25 and 1984 X.400 (an ATS for 1984 was published with the 1988 X.400 recommendations).

Facing a dearth of standard ATSs, we issued a public call for ATSs to meet GOSIP requirements, specifically: cover the necessary functions in appropriate breadth and depth, exist in the public domain free of licensing charges and copyrights, and, if possible, be specified in the standard test language (i.e., tree and tabular combined notation). We received responses from the Corporation for Open Systems (COS) in the United States and from the Opens Systems Testing Consortium (OSTC) in Europe. Many of the test descriptions were not in the accepted standard test language, but we did not disqualify them. The OSTC submission, initially under copyright, was placed into the public domain for our intended purpose. The COS submission was not under copyright. The most difficult problem was evaluation of the coverage of the ATSs. To perform the review we asked the help of the OIW, which changed its charter in September 1989 to include the review and refinement of abstract tests. The review meetings commenced in January and were held monthly through May 1990. Although the reviewers included experts from user organizations, OSI suppliers, test system developers, and NIST, the volume of material to review, the detailed technical nature of the work, and the rigid time constraints combined to create a

difficult task. We are not totally pleased with the thoroughness of the reviews and the resulting ATSs are deficient in some areas, notably FTAM, but we are prepared to publish and distribute them and to use them as the basis for GOSIP Version 1.0 testing. With the help of others, we expect to add incremental improvements, over time, to the ATSs and to add ATSs to match new functionality in subsequent versions of GOSIP.

Means Of Testing

Agreed ATSs provide a basis to solicit and assess test systems, so-called means of testing (MOTs). A MOT is a combination of hardware and software that can execute tests, in cooperation with a vendor's product, and evaluate the result of each test (pass, fail, or inconclusive). Development of MOTs for OSI protocol conformance is a relatively immature field; the first such testers were developed for the transport protocol in about 1982. Because MOT development is immature and because a number of MOT suppliers exist, we concluded that competition among MOT suppliers could lead to rapid improvements in the quality of MOTs; thus, we decided to solicit multiple candidate MOTs, to evaluate each candidate on its merits, and to permit any and all MOTs that satisfied our criteria to be used in the GOSIP testing program. The following discussion permits only the briefest outline of the procedures and criteria; more detail is provided in the *MOT Assessment Handbook* [3].

Recognizing that the candidate set of MOTs might benefit from substantial improvements, we established three levels of merit within the GOSIP test program: adequate, provisionally acceptable, and unacceptable. To be judged adequate, an MOT must implement the appropriate ATS in the required breadth and depth, must possess sufficient quality and maintainability as demonstrated by the software and documentation, must be useable by a competently trained test operator, and must be available for license to multiple sources. To be judged provisionally acceptable, an MOT must meet a defined level for each of the criteria, a level that is deemed acceptable, but not fully adequate. We felt such compromise was necessary at the present state of MOT development. When at least one MOT for a given ATS is judged adequate, no provisionally acceptable MOTs will be qualified; otherwise, all provisionally acceptable MOTs for a given ATS will be qualified. MOTs not meeting at least the criteria established for provisional acceptance will not be qualified.

MOT assessment requires both static and dynamic analysis. Static analysis entails comparing the MOT capability against the appropriate ATS, protocol conformance requirements, and MOT criteria, as maintained by NIST. Dynamic analysis requires executing a set of tests, selected by the MOT assessor, against a known sample (that is, an implementation for which a known set of test results is expected) selected by NIST. MOT assessment procedures allow for acceptance of MOTs after only a static analysis, if the MOTs are derived (e.g., moved to a new hardware platform), and assuming that the original MOT had successfully passed dynamic analysis.

NIST exercised the MOT assessment procedures during a cooperative research and development agreement with the Corporation for Open Systems (COS). Between January and August 1990, NIST and COS worked together on a trial assessment of COS testers, enabling NIST to refine the proposed procedures and helping COS improve the capabilities of several MOTs. Between April and May 1990, NIST announced an informal call for potential MOT suppliers. Each responding MOT supplier received the *MOT Assessment Handbook.*

Accredited Test Laboratories

After an ATS is approved and a method of qualifying MOTs is established, GOSIP testing laboratories can be accredited through the National Voluntary Laboratory Accreditation Program (NVLAP), using technical and administrative procedures constructed jointly by the National Computer Systems Laboratory (NCSL) and NVLAP. (NCSL and NVLAP are organizational units within NIST.) The detailed laboratory accreditation procedures are described in the *GOSIP Testing Laboratory Accreditation Handbook* [4]; here we recount only a few of the concepts.

NVLAP provides an independent means of accrediting either first (vendor) or third (contracted) party laboratories for testing for conformance to GOSIP. The candidate laboratories must use MOTs previously qualified by NCSL, or its agent, and must demonstrate competence in using the MOTs to conduct test campaigns. Accreditation usually involves a site visit by one or more independent assessors selected by NVLAP.

During August 1990, NCSL and NVLAP conducted a pilot accreditation using laboratory facilities at COS. In May and June 1990, NIST announced an informal call for potential GOSIP testing laboratories. Each responding candidate received the *GOSIP Testing Laboratory Accreditation Handbook* and a list of potential MOT suppliers.

Interoperability Testing

Although necessary to assess adherence to a standard and to detect software errors in protocol implementations, conformance testing does not demonstrate interoperability among OSI products; and yet, interoperability, a key goal of OSI, is of utmost concern to users. Thus, while there is no technical consensus about its role, interoperability testing is a necessary part of any OSI testing program.

For the GOSIP Testing Program two types of interoperability testing, both based on a single, consensus set of interoperability tests, are recognized: testing of a supplier product with a GOSIP Reference Implementation and bi-lateral testing of pairs of supplier products. The consensus interoperability tests currently are the FTAM and X.400 tests originated by OSINET, an informal group of OSI suppliers and users, and subsequently coordinated among other similar groups aggregated under the name OSIone. The tests, to be published by NIST and registered for use in GOSIP testing, are within the public domain. A subset of the OSINET tests has been selected by NIST for verifying the interoperability of GOSIP-conformant products. The requirement for a supplier to demonstrate interoperability with a GOSIP Reference Implementation (if, and only if, such implementation is registered by NIST) is debatable. The many valid arguments, both pro and con, are too involved to recount here. We believe that interoperability against a known reference implementation will yield increased product quality, provide a check on the conformance testing process, and aid evaluation of interoperability test suites.

While these same benefits may accrue via bi-lateral testing

among pairs of OSI product suppliers, we are unable to formulate a sensible and fair statement of policy. Thus, we are encouraging industry cooperative efforts to build voluntary services for the registration of results from successful bi-lateral interoperation testing.

NIST will evaluate candidate OSI interoperability testing services. Successful candidates, to be identified on a publicly accessible register, must:

1. be an organization recognized by NIST,
2. use an interoperability test suite recognized by NIST,
3. arrange for a bi-lateral test agreement between pairs of GOSIP product suppliers,
4. select a common subset of tests including the mandatory tests for GOSIP,
5. provide a joint declaration from each pair of test partners for each successful test campaign, and
6. make available, upon request, to NIST a copy of the detailed test results for any specific test campaign.

Government users experiencing interoperability failures among pairs of OSI products for which successful interoperability results are documented through a testing service registered with NIST can, after exhausting their appeals with the appropriate vendors and testing service, request assistance from NIST. After investigating the problem, NIST can, if warranted, request that the testing service remove the suppliers declaration of interoperation for the product pairing. Further, NIST can, if warranted, ask that the testing service arrange for NIST to witness interoperability retests for any extant suppliers declaration of interoperation involving either of the two products in question. Should the testing service reject these requests, then NIST can withdraw recognition. These procedures for ultimate appeal to NIST should increase confidence among government users and encourage existing industry initiatives for OSI interoperability testing.

Public Registers

The key to the GOSIP Testing Program is a set of publicly accessible registers maintained by NIST. Details concerning these registers are given in the *GOSIP Testing Registration Criteria* document [5]. Management of the GOSIP Testing Program is complicated by change: implementor agreements, Abstract Test Suites, Means of Testing, and GOSIP products are all evolving. To monitor and direct staged improvements in the deployed base of GOSIP products, NIST has established registers for:

1. GOSIP Abstract Test Suites,
2. Interoperability Test Suites,
3. Assessed Means of Testing,
4. Accredited Test Laboratories,
5. GOSIP Reference Implementations,
6. Conformance Tested GOSIP Products,
7. Interoperable GOSIP Products, and
8. Interoperability Testing Services.

Today the Abstract Test Suites, incomplete and under-evaluated, are provisionally registered, subject to staged improvements on a yearly basis. The Means of Testing and GOSIP products which depend on the ATS will also be provisionally registered. As the ATSs reach stability - probably as International Standard ATSs are achieved - and as MOTs and GOSIP products mature, the provisional nature of registration will be dropped and, ultimately, as the pace of change in OSI abates, the need for registration could cease.

Schedule

Our goal is to accredit an initial group of GOSIP testing facilities by November 15, 1990, leading to availability of the first demonstrably GOSIP conformant products as early as December 1990 or January 1991. The first milestone was July 15, 1990 when we issued a formal call for candidate MOTs and proposed GOSIP testing laboratories. The call closes August 15, 1990. From August 15 to October 15 we are assessing MOTs. From September 15 to November 15 we are evaluating laboratories. It is possible that, in some instances, successful product testing will have been previously conducted in laboratories accredited subsequently, and, therefore, some GOSIP products may be 'grandfathered' onto the register by December 1. This schedule is tight.

Testing Policy

Our policy for applicability of the GOSIP testing criteria is crafted carefully to satisfy two competing needs: the need to reduce the cost of testing and the need to provide market confidence in suppliers claims of GOSIP conformance and interoperability. The GOSIP FIPS 146 requires Federal Government Agencies, when acquiring computer network products and services, to procure OSI products, as specified in GOSIP. A companion FIPS, *GOSIP Conformance and Interoperation Testing and Registration*, places certain responsibilities on GOSIP product suppliers and makes recommendations to Acquisition Authorities. A summary follows.

1. If a supplier claims GOSIP conformance for a product, then that product must be tested in accordance with the criteria specified in the *GOSIP Conformance and Interoperation Testing and Registration* FIPS. If the product includes a multi-layered OSI profile, then **all** protocols for which GOSIP conformance is claimed must be tested in accordance with the FIPS criteria.
2. Federal Government Agencies requiring validation of supplier claims of GOSIP conformance should consult the register of Conformance Tested GOSIP Products.
3. Federal Government Agencies wishing to procure OSI products that are not on the register of Conformance Tested GOSIP Products are advised to arrange that the product qualify for the register prior to accepting the product or to stipulate contractually that the supplier shall arrange that the product qualify for the register by a specific date.

4. Federal Government Agencies requiring an increased confidence that a specific GOSIP-conformant product will interoperate should consult the register of Interoperable GOSIP Products, **but only if** a corresponding GOSIP Reference Implementation is registered **and** that registration is not provisional.
5. Federal Government Agencies should consult the data supplied by a registered Interoperability Testing Service under any of three conditions:
   (a) the Agency requires increased confidence that a specific GOSIP-conformant product will interoperate, but no corresponding GOSIP Reference Implementation is registered or a corresponding GOSIP Reference Implementation is only provisionally registered,
   (b) the Agency requires that multiple instances of successful interoperation are documented for a specific GOSIP-conformant product, or
   (c) the Agency requires that an instance of successful interoperation is documented for one or more specific pairs of GOSIP-conformant products.

Assessment

We believe our efforts have focused U.S. testing requirements in a structure useful for the near-term, say the next five years. We will have established an operational system for endorsing ATSs, for qualifying MOTs, for accrediting test laboratories, and for registering GOSIP-conforming products. The system is new, untested, and subject to evolution, but we have a start. The first set of ATSs have not been reviewed as thoroughly as we had hoped. The known FTAM test suites do not include tests for the T2, M1, and A2 profiles, and no public domain source of IEEE 802.5 tests is known. The MOTs available today are not stable and well tested, nor do they comply with every detail of GOSIP. We expect the number and competence of GOSIP testing laboratories to grow over the next few years. In summary, we started very late, perhaps lulled to sleep by numerous industry initiatives in the U.S., Europe, and Japan which have not achieved the early success expected.

Looking Ahead

The largest challenge barring our path to open systems is the need to manage change. Changes to the base standards, implementor agreements, procurement specifications, products, tests, and test systems must be synchronized. Major inhibitors to such synchronization abound. For most base standards a companion set of standard tests are not defined. Implementor agreements are not yet coordinated world-wide and, of course, no globally agreed set of tests exists. A large number of procurement specifications exist world-wide: a good indication of market potential, but a possible source of market fragmentation. Remarkably, interoperable OSI products exist from a significant set of computer vendors.

Moving ahead will take a coordinated, global investment from a variety of sources: users, vendors, and governments, working together to create a common good. The challenge is to direct our investment for the highest productivity and the largest return. How can we answer the challenge?

The necessity of internationally agreed standard profiles is now clear. With three regions of the world making implementation agreements and with government sector procurements throughout the world requiring OSI, the penalty for unnecessary divergence is a market fragmentation that bodes ill for the prospects of international interoperability and for the potential of an integrated, open world market for information technology (IT) products and services. The jobs of product development, test system creation, and operational deployment become more difficult, more expensive, and less beneficial, as the diversity among standards increases.

The benefits of uniform testing requirements are also becoming apparent. If a vendor can build a product, test it once, and then have the product and test results accepted throughout the world, the cost of product development will be significantly reduced. Thus, we need to produce a set of tests, test methods, and testing procedures that might be accepted around the world. We at NIST believe the first window of opportunity for aligning OSI testing requirements between the U.S. and Europe will come around the end of 1991, when we must have a testing program in place for U.S. GOSIP Version 2.0. If that date is too ambitious, another opportunity will appear as GOSIP Version 3.0 is prepared. Our message is that an alignment must occur, if we are to achieve the benefits of uniform testing requirements.

Finally, the buyers of the world must, to the extent possible, agree on a set of procurement specifications, including testing requirements, that are consolidated, giving IT product vendors a large incentive to hit the mark. A fragmented market weakens the case for an investment in standard products and encourages competition among powerful interests to set de facto IT standards. Where de facto standards rule, the buyers invest in such standards through product purchases, trusting the dominant vendor to manage the evolution of the standards. While reliance on a dominant supplier has a price, how can we persuade users to invest in the public standards process, if we cannot make a convincing case that we are prepared to manage change?

As a final thought, looking ahead, we must remind ourselves that, even having met the technical and management challenges, international trade and integrated economies are entangled in a web of political and economic considerations that sometimes, rightly or wrongly, supersede other issues, and obstruct our way forward. We must carry our message to the policy makers in corporations and governments. Technical solutions must be in place, if our message is to be clear and convincing.

References

[1] *The Government Open Systems Interconnection Profile.* FIPS PUB 146, U.S. Department of Commerce, Washington, D.C., August 1988.

[2] J. Stephen Nightingale. *GOSIP Conformance and Interoperation Testing and Registration.* August 1990. Available from the author at: NIST, Building 225 Room B217, Gaithersburg, MD 20899 USA.

[3] J. Stephen Nightingale. *Means of Testing Assessment Handbook.* August 1990. Available from the author at: NIST, Building 225 Room B217, Gaithersburg, MD 20899 USA.

[4] J. Stephen Nightingale. *GOSIP Testing Laboratory Accreditation Handbook.* August 1990. Available from the author at: NIST, Building 225 Room B217, Gaithersburg, MD 20899 USA.

[5] J. Stephen Nightingale. *GOSIP Registration Criteria Handbook.* August 1990. Available from the author at: NIST, Building 225 Room B217, Gaithersburg, MD 20899 USA.