Towards a Trustable Virtual Organisation

Jun Ho Huh and Andrew Martin
Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

Introduction

A wide range of research is conducted, archived, and reported in the digital economy. Its influence has grown over the years to include various disciplines from science through to finance and industrial engineering. In consequence, different types of distributed systems have been deployed to facilitate the collection, modelling and analysis of the dispersed data; or the sharing of the computational power.

A problem arises, however, when the models or data contain sensitive information or have commercial value. These then become lucrative targets for attack, and may be copied or modified by adversaries. This is particularly true in many of the scientific disciplines where researchers – who care about the *confidentiality* of their sensitive data or the *integrity* of the collected results – are reluctant to exploit the full benefits of distributed computing. Submitting a privileged job to the distributed resources requires prior knowledge of the security standards of all of the target systems — only those running with acceptable security configurations and patch levels should be selected to execute the job. However, this still remains as a 'trust gap' between the job owners' requirements and current technological capabilities, and serves as a barrier to up-take of existing systems.

For instance, a Condor system [1] may provide a digital certificate infrastructure for users to identify resource/service providers and vice versa. Without robust mechanisms to safeguard the keys from theft, however, this solution offers only a modest improvement over legacy architectures. Moreover, rogue administrators might replace the compute nodes with malicious ones, tamper with them, or subvert their security configurations to steal sensitive data and/or return fabricated results.

Trusted Computing

Faced with the prospect of modern PCs (and other devices) having so much software that their behaviour is unpredictable and easily subverted, the Trusted Computing Group (TCG) has developed a series of technologies based around a Trusted Platform Module (TPM) which helps to provide two novel capabilities [2]: a cryptographically strong identity and reporting mechanism for the platform, and a means to measure reliably a hash of the software loaded and run on the platform (from the BIOS upwards). These provide the means to *seal* (encrypt) data so that it will only successfully decrypt when the platform measurements are in a particular state; and to undertake *remote attestation*, proving to a third party that (in the absence of hardware tampering) a remote device is in a particular software state.

Trusted Computing Approaches

Great strides have been made in using trusted computing to design and construct trustworthy components for distributed systems. For example, as a possible solution to the 'malicious host' problem, Cooper and Martin [3] describe a grid architecture that ensures protected job execution environments. Remote attestation allows users to verify that this environment is in place before submitting their jobs. Löhr et al [4] propose a 'Trusted Grid Architecture', in which users collect configuration tokens (platform-configuration-based credentials) of service providers, and verify their security configurations using a locally managed application whitelist.

These approaches, however, have shortcomings in the areas of platform configuration management and provision of *usable* attestation services. The users are often expected to manage application whitelists, and evaluate the security properties of remote systems. This is unrealistic because: (1) the whitelist entries will require constant modification and update, and (2) normal users (e.g. researchers) will not have sufficient security knowledge (and experience) to evaluate such properties and make *trust decisions* on their own.

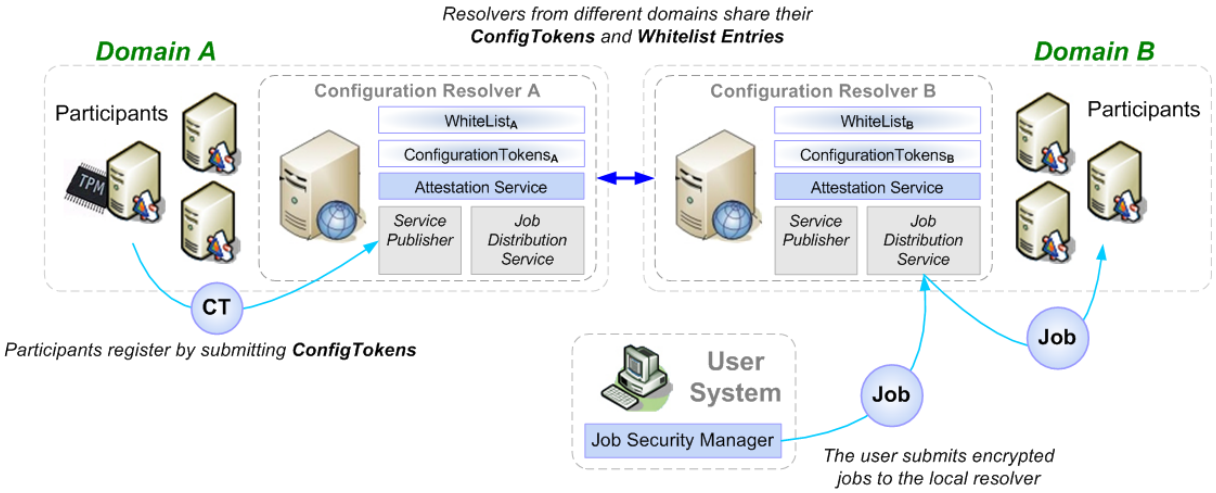

Figure 1 The configuration resolver is added to each domain to manage an up-to-date directory of trustworthy participants (ConfigTokens) and handle job distribution process.

To bridge the 'trust gap', we propose two different types of trustworthy distributed systems – one applicable for a computational system and the other for a distributed data system. Central to the distributed systems is the novel idea of the 'configuration resolver'. In both designs, the configuration resolver maintains an up-to-date whitelist and performs attestation on the user's behalf, ensuring that the jobs are dispatched to only those considered trustworthy (see Figure 1). This aims to provide a more usable attestation service for large-scale distributed systems.

The job secrets (e.g. models and data) are encrypted with the public key of the trustworthy participant, and safely distributed over the unprotected network. The private half will only be accessible if the security configurations of the participant's trusted computing base and the job virtual machine remain unchanged. Runtime verification of the virtual machine integrity guarantees a trustworthy execution environment.

In the distributed data system, the configuration resolver operates inside the 'blind analysis server', and together, they provide a trustworthy environment to run statistical analyses on the raw data without disclosing it to anyone. Attestation of the blind analysis server is sufficient to establish that only the processed, anonymised results will be released to the user. The main advantages are that no information is lost through anonymisation (prior to release of data), and in consequence, analyses are carried out on more accurate datasets, producing high-quality results.

Conclusion

We describe two security architectures for distributed systems based on trusted computing capabilities. In both architectures, a central 'configuration verification server' ensures that jobs are dispatched to trustworthy participants and executed in protected environments. As a form of evaluation, we suggest how the proposed ideas could be integrated with existing grid/cloud systems, and highlight the potential security enhancements.

References

[1] D. C. H. Wallom and A. E. Trefethen, "Oxgrid, a campus grid for the university of oxford," *in UK e-Science All Hands Meeting*, 2006.

[2] D. Grawrock, The Intel Safer Computing Initiative. Intel Press, 2006, pp. 3–31.

[3] A. Cooper, A. Martin, Towards a secure, tamper-proof grid platform, *CCGRID 06. Sixth IEEE International Symposium on*, 2006.

[4] H. Löhr, H. V. Ramasamy, A.-R. Sadeghi, S. Schulz, M. Schunter, C. Stuble, Enhancing grid security using trusted virtualization, *in: Autonomic and Trusted Computing*, Springer 2007.