Security for Data Intensive Computing

Kevin L. Wohlever

OSC - Springfield

OSC
SPRINGFIELD

Data Security Thoughts

- **"The need for data security is in direct conflict with the requirement for increased accessibility and sharing of data banks. The only certainty in this rapidly developing field of data security appears to be that technical solutions to avert known threats will be developed. With equal certainty, new threats will evolve. To paraphrase a famous saying: eternal vigilance is the price of data security."**

- **IEEE Computer Magazine (p31)- February**

- **1974**

Question 1

- Are there differences in security requirements for transient data usage within a computational cluster vs. longer term storage? If so, what are the differences? Do they lead to opportunities to apply stronger security mechanisms to longer term storage without getting in the way of high-performance computing usage of that data? If so, how?

Responses to Q1

- Are there differences in security requirements for transient data usage within a computational cluster vs longer term storage? If so, what are the differences?
  - Yes there are differences in security requirements for transient data vs longer term storage.
  - Differences will need to be identified individually. This is a risk analysis issue.

More Question 1 Responses

- Do they lead to opportunities to apply stronger security mechanisms to longer term storage without getting in the way of high-performance computing usage of that data? If so, how?
  - For example, use the excess capacity of storage controllers to do encryption / decryption
  - But this raises more questions

Responses to Q2

A) DoD Orange book requirements were accepted into HPC computing by a number of OS / System vendors (Cray, SGI, Sun)

B) Commodity OS environments are the key gaps. Until the majority of their users, personal or enterprise, require additional security, it will not be implemented.

C) We haven't been able to address gaps in the past 30 years, what makes us think we can do it now?

Responses to Q3

A) Data needs to be secured and tracked

B) ACLs, access logs should become more prevalent

C) What are the opportunities???

  A) Individual security requirements, vs system wide

  B) automated encryption

  C) public key issues are raised

  D) Accounting and charging improvements may help

Question 4

- Given commercial (open source) security developments in these and other areas, what are the most important gaps that need to be addressed to meet the security needs of high performance computing sites? What can/should the high performance computing site do to motivate more, or more in depth, effort to address these gaps?

WARNING

The Following is a Pseudo Self Serving Promotion

Background

DICE

Grand "Experiment" to see if…

- **Different government agencies can:**
  - **Develop a partnership with the HPC vendor community?**
  - **Partnership will conduct creditable evaluations of emerging technology solutions?**
- **Intellectual Property can be protected?**
- **The *"Time to Solutions"* of complex problems can be improved?**

ey Individuals

Agency Representatives

- DoD – Mr. Steven Wourms
- NASA – Dr. Phil Webster
- DOE – Dr. Neil Pundit
- OSC – Dr. Stan Ahalt

Test Site Representatives

- DoD – Mr. Lloyd Slonaker
- NASA – Dr. Dan Duffy
- DOE – Mr. Lee Ward
- OSC – Mr. Kevin Wohlever

DICE Management Team

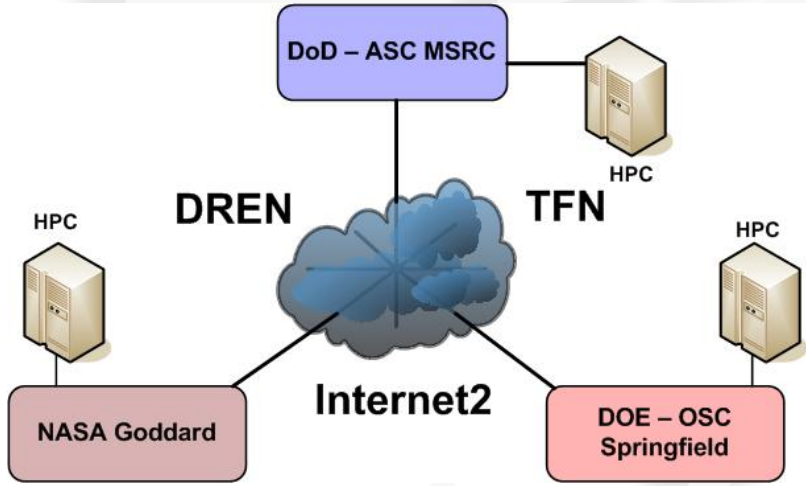
- Principal Investigator – Mr. Rob Evans-Miller, AVETEC
- Program Manager – Mr. Roger Panton, The Greentree Group
- Program Manager/Integrator – Mr. Tracey Wilson, CSC

Objectives and Goals

- **DICE Objective**
  - **Evaluate current/emerging data management technologies for their ability to improve *data accessibility* over geographically distributed sites**
- **DICE Goals**
  - **Promote Government/Supplier/Research Collaboration**
  - **Establish a distributed collaborative inter-agency test environment**
  - **Promote solutions to issues such as: data locality, movement of data, data security and data manipulation**

Test Bed Requirements

1. *Geographical separation for data intensive evaluation.*
2. *Each location connected over existing high-speed networks with HPC, data storage, and application SW.*
3. *Data must be public releasable*
4. *Project "outputs" must be free-standing and deployable within the respective agencies.*
5. *Minimize connectivity with "Centers" production HPC, network and storage resources*



Benefits

Agency

- Implement tested and evaluated solutions
- Positive feedback between agency and the vendor
- Provides “One Voice” direction to the vendor
- Real world evaluation of new technologies and capabilities
- Establish industry standard metrics/benchmarks

Vendor

- “One Voice” focuses efforts on advancements that address customer needs
- Creates real-world environment to test emerging technologies and advanced features
- Creates better understanding and dialog with customers
- Provides multi-agency reference source for new products

Next Step

- **Initial test bed operation**
  - **Setup individual sites end of March 2006**
  - **Operations between all three sites April 2006**
  - **In-house evaluations conducted April – May 2006**
- **Project schedule**
  - **Call for projects March 6, 2006**
  - **Project submitted by end of April 28, 2006**
  - **Technical Review Board evaluation complete May 26, 2006**
  - **Governance Board approval May 31, 2006**
  - **First project evaluation started June 2006**