

Data Integrity – Is Your Data the Data You Really Think It Is????

Byron V. Peters

Raytheon Company

1616 McCormick Drive, Upper Marlboro, MD 22152

Phone: +1-301-925-0350 FAX: +1-301-925-0326

E-mail: bpeters@east.hitc.com

Presented at the THIC Meeting at the Raytheon ITS Auditorium

1616 McCormick Drive, Upper Marlboro MD 20774-5301

October 26-27, 2004

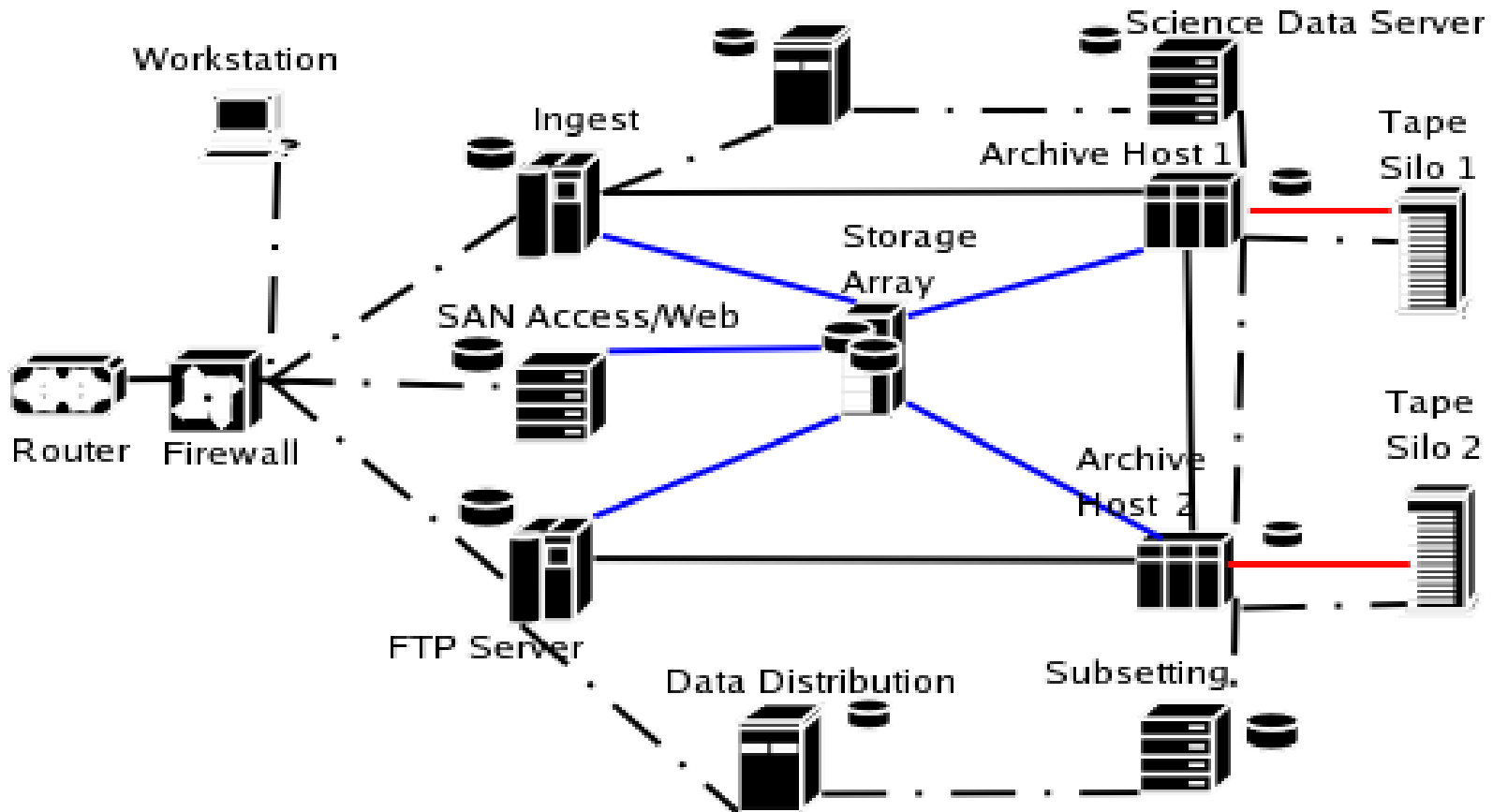
A Little Nite Flying...



Photo Copyright © Freight-Dawg

AIRLINERS.NET

...To a Little Nite Fright



Know your Equipment...

- Complexity bad – Simplicity good
- Reliance on Pre Operation Status Tests (POST)
 - Almost all components do an internal check-goodness – but not complete
- The insidious “dumb” interface connector(s) - cable – connector(s)
- Even with decreasing chip count, “dumb” interfaces abound

Data Integrity Alternatives

- Hardware
 - End-to-End hardware verification
- Software
 - Checksum-based verification

“Smart” Interfaces

- Microprocessors are everywhere
 - Where there is a microprocessor, a pre-operation test is almost always available
- POSTs test the internal operation from in to out without checking the input or the output

“Dumb” Interfaces

- We need a Star Trek
 “Level 1 Diagnostic”!
 - Not yet even conceived...
 - SNMP the closest thing
 - Lots of cables left inside of systems
 - No method of verifying connector to connector soundness
 - Potential problem any time the wires are disturbed

Assume Nothing...

- Checksum and track every data file
 - On ingest, get checksum from provider
 - On processing, verify output and generate checksum

Know Your Network

- “Not well known” weakness in Ipv4
 - 16-bit checksum – protocol designed when 1TB was rare...
 - Can NOT tell the difference between 0x0000 and 0xFFFF
- Numerous Network Interfaces – Many of the “dumb” interfaces belong to the network
 - Ethernet
 - SCSI
 - HBA

Where to From Here?

- Hardware
 - Give me a level 1 diagnostic!
- Software
 - Improve high-level applications with better checksum awareness
 - Design and implement an
END-TO-END
verification capability in software

Summary

- Enterprise-class systems have an inherent data integrity weakness due to their complexity
- The more complex the system, the more imperative is a checksum management system
- Hardware and/or software-based end-to-end system verification would be a major step in improving the trustworthiness of data-intensive systems