Grand Challenge:

Dependability Benchmarking & Prediction

(IFIP WG 10.4 Benchmark SIG)
http://www.dependability.org

Chair: Phil Koopman, Carnegie Mellon University
Vice Chair: Henrique Madeira, University of Coimbra

First Meeting: November 1st, 1999
Problem Scope

- **Implementation Technology**
  - Hardware, software, control algorithms, user interface, mechanical safety backups

- **Operational life cycle**
  - Specification, design, deployment, maintenance, operation, disposal

- **Product deployment scale**
  - Capital equipment, business infrastructure, consumer products, disposable goods

- **What are the stakes?**
  - Would you bet your **life** on a computer running off-the-shelf software?

  You will...
Grand Challenge Goal

◆ Be able to *predict* the dependability of a critical system before first product shipment
  • Including all real-world issues
  • Both comparative and absolute metrics if possible

◆ Major Issues:
  • Can we subset the problem space for tractability … and still be useful?
    – Real Time Mission Critical Systems offer an attractive starting point
  • Can we capture all the critical tradeoffs in any single “benchmark number”?
    – Too many numbers is confusing; too few might be overly simplistic
    – And all the other usual problems with benchmarking
What Makes Dependability Challenging?

- **Dependability prediction for electronic hardware exists**
  - A result of World War 2 adoption of electronics

- **Existing approaches**
  - Brute force redundancy is OK for hardware, but expensive
  - Many approaches assume a perfect design/specification/etc.
    - Software fault tolerance is still an evolving field

- **We still aren’t very good at accounting for:**
  - Software
  - People
  - QoS and “soft” dependability issues
  - Security (“malicious” faults)
  - Systems with imperfect maintenance/support/…
  - Systems with constrained budgets
Possible Elements of a Benchmark

Create a benchmarking scenario:

• Specifications of expected system behavior in different fault situations
• An operating scenario with a workload
• A faultload, used to inject:
  – System faults, exceptional situations, component overloads, operator mistakes, maintenance errors, component failures, etc.
• Procedures and rules for benchmarking activities
• Instrumentation to record the above
• Measures based on instrumentation
Possible Alternate Approaches

- Based on discussions in first meeting
  - Important to get numbers; but they must mean something useful
  - Benchmarking might be too aggressive for a first attempt

- “Consumer Reports” Approach
  - Measure whatever we can measure
    e.g. using fault injection; historical trends
  - Weave a pattern about the system based on this information

- Piggyback on an existing benchmark
  - TPC/C + dependability?

- Use a process-based approach
  - SEI CMM plus dependability best practice?
IFIP WG 10.4 Benchmarking SIG Goals

- **Exchange of ideas**
  - Promote cross-pollination; reduce wasted effort

- **Document state of the art**
  - Set of research white papers on what seems to work; what doesn’t *(e.g., “Grey Series” dependability book from DCCA series)*
  - What can we borrow from others; what is unique to dependability?
  - Encourage/publicize existing tools & techniques

- **Create issues list**
  - What issues must a dependability benchmark address?
  - What about multidimensional composites (performance, security, *etc.*)
  - What are the constraints that must be faced to attain success?

- **Propose path to dependability benchmarks**
  - Or at least published position papers on the alternatives

- **Next meeting in San Jose, early April 2000**
  - Contact: koopman@cmu.edu