# Final Report for CCC Cross-Layer Reliability Visioning Study: Executive Summary March 3, 2011

The geometric rate of improvement of transistor size and integrated circuit performance known as Moore's Law has been an engine of growth for our economy, enabling new products and services, creating new value and wealth, increasing safety, and removing menial tasks from our daily lives. Affordable, highly integrated systems have enabled both life-saving technologies and rich entertainment applications. Anti-lock brakes, insulin monitors, and GPS-enabled emergency response systems save lives. Cell phones, internet appliances, virtual worlds, realistic video games, and mp3 players enrich our lives and connect us together. Over the past 40 years of silicon scaling, the increasing capabilities of inexpensive computation have transformed our society through automation and ubiquitous communications.

Looking forward, **increasing unpredictability threatens our ability to continue scaling integrated circuits at Moore's Law rates.** As the transistors and wires that make up integrated circuits become smaller, they display both greater differences in behavior among devices designed to be identical and greater vulnerability to transient and permanent faults. Conventional design techniques expend energy to tolerate this unpredictability by adding safety margins to a circuit's operating voltage, clock frequency or charge stored per bit. However, the rising energy costs needed to compensate for increasing unpredictability are rapidly becoming unacceptable in today's environment where power consumption is often the limiting factor on integrated circuit performance and energy efficiency is a national concern. Reliability and energy consumption are both reaching key inflection points that, together, threaten to reduce or end the benefits of feature size reduction.

To continue beneficial scaling, we must use a *cross-layer, full-system-design approach to reliability*. Unlike current systems, which charge every device a substantial energy tax in order to guarantee correct operation in spite of rare events, such as one high-threshold transistor in a billion or one erroneous gate evaluation in an hour of computation, cross-layer reliability schemes make *reliability management a cooperative effort across the system stack*, sharing information across layers so that they only expend energy on reliability when an error actually occurs. Figure 1 illustrates an example of such a system that uses a combination of information from the application and cheap architecture-level techniques to detect errors. When an error occurs, mechanisms at higher levels in the stack correct the error, efficiently delivering correct operation to the user in spite of errors at the device or circuit levels.

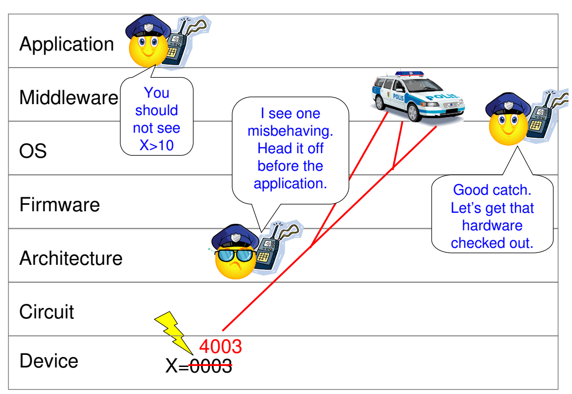


Figure 1: Cross-Layer Cooperation

In the realms of memory and communication, engineers have a long history of success in tolerating unpredictable effects such as fabrication variability, transient upsets, and lifetime wear using information sharing, limited redundancy, and cross-layer approaches that anticipate, accommodate, and suppress errors. Networks use a combination of hardware and software to guarantee end-to-end correctness. Error-detection and correction codes use additional information to correct the most common errors, single-bit transmission errors. When errors occur that cannot be corrected by these codes, the network protocol requests re-transmission of one or more packets until the correct data is received. Similarly, computer memory systems exploit a cross-layer division of labor to achieve high performance with modest hardware. Rather than demanding that hardware alone provide the virtual memory abstraction, software page-fault and TLB-miss handlers allow a modest piece of hardware, the TLB, to handle the common-case operations on a cycle-by-cycle basis while infrequent misses are handled in system software.

Unfortunately, mitigating logic errors is not as simple or as well researched as memory or communication systems. This lack of understanding has led to very expensive solutions. For example, triple-modular redundancy masks errors by triplicating computations in either time or area. This mitigation methods imposes a 200% increase in energy consumption for *every* operation, not just the uncommon failure cases.

*At a time when computation is rapidly becoming part of our critical civilian and military infrastructure and decreasing costs for computation are fueling our economy and our well being, we cannot afford increasingly unreliable electronics or a stagnation in capabilities per dollar, watt, or cubic meter.* If researchers are able to develop techniques that tolerate the growing unpredictability of silicon devices, Moore's Law scaling should continue until at least 2022. During this 12-year time period, transistors, which are the building blocks of electronic components, will scale their dimensions (*feature sizes*) from 45nm to 4.5nm. This additional scaling will be possible only if we can mitigate unpredictability and noise effects in small feature-size devices as effectively as we have been able to compensate for these effects in memory—that is, without paying an increasing energy overhead. The challenge for the near future, then, is to achieve the density and full energy benefits of ideally-scaled smaller technologies with the reliability of our larger and older technologies. Over the longer term, techniques to tolerate unpredictable behavior at low energy costs will also be necessary to make post-silicon technologies, such as quantum and molecular computation, feasible.

Several corporations shared their perspective on how reliability affects them economically:

*"Microprocessor and computer system reliability is a critical issue for Intel. As we look forward, we are deeply concerned about the techniques used to tolerate errors, device variations, and silicon aging and their impact on the cost, performance, and power consumption of our products. Meeting the reliability challenges of future products will require innovative new approaches in which the entire system contributes to overall reliability, and we strongly endorse research into these architectures and methods."* – Justin Rattner, Vice-President and Chief Technology Officer, Intel Corporation.

*"Reliability and fault tolerance are essential to the performance of many embedded systems. For example, automotive, industrial and medical applications may involve safety-critical functionality and harsh operating environments. As the scope and complexity of these applications continue to increase, new approaches are needed to design, validate and qualify highly reliable and resilient embedded systems."* – Ken Hansen, Vice President and Chief Technology Officer, Freescale Semiconductor Inc.

*"Achieving high reliability in the types of high end systems that IBM makes is a very high priority for us. We invest significant design and engineering resources into the hardware, firmware and software aspects of error detection, correction, and avoidance. As the potential for these errors increases due to technology scaling, additional investments must be made, and new ideas proposed to find solutions that are cost effective, practical and achievable within current design paradigms. This is not a problem that we expect to go away anytime soon, and solving it must have a high priority."* – Carl J. Anderson, IBM Fellow.

To better understand the reliability challenges facing future electronic systems, the limitations of current approaches, and the opportunities offered by cross-layer techniques that distribute the responsibility for reliability across the entire system stack, the Computing Community Consortium funded a study into cross-layer reliability, which was carried out from March-October of 2009. All told, eighty people participated in one or more of the study group's three meetings, representing academia, industry, and government.[1] During the study, we formulated constituency groups on commercial and consumer electronics, infrastructure, life-critical, space/avionics, and large-scale systems. These groups covered computing systems ranging from pacemakers and hand-held computers to satellites and supercomputers, as well as applications ranging from entertainment to protecting human life. Each of the groups had different perspectives on how scaling-induced power and reliability challenges impacted them and were willing to make different tradeoffs to address those challenges. Nonetheless, all of the industries represented are feeling pain from these reliability problems and agree that the coming challenges will not be satisfied by current solutions.

Throughout the study, participants described *situations where the* **lack of global information** *caused designers to make worst-case assumptions at individual design layers, resulting in systems that were over-designed, inefficient, and overly expensive.* Looking forward, many of the participants identified cases where current design techniques would be unable to meet the needs of future systems, making a new approach to reliable system design necessary instead of merely desirable.
**Common challenges that should be addressed include:**

- *Late-Bound Information:* Critical information about how a system or component will be used and/or the characteristics of the technology that will be used to manufacture it is often unavailable until late in the design process. Inflexible system designs cannot adapt to this information as it becomes available, resulting in over-design, limited ability to use the system in different contexts, and occasional need for complete re-design late in the design process.
- *Lack of Instantaneous Operational Information:* Error rates in electronic systems vary substantially with operating conditions, such as temperature and altitude. Systems that cannot sense and react to changes in their environment must always inefficiently assume worst-case conditions.
- *Lack of Information on Application Requirements:* Different applications (mp3 player vs. power plant control) have very different reliability requirements. Hardware that does not know an application's needs must provide complete reliability even if the application can tolerate errors.
- *Lack of Information on Health of Components* and *Lack of Information on Components from Heterogeneous Suppliers:* Electronic components vary from manufacturer to manufacturer, from part to part from the same manufacturer, and from time to time on the same part. Systems that cannot analyze and adapt to these differences must assume worst-case behavior, leading to designs with poor performance and high power consumption. Software that is

[1] See `http://www.relxlayer.org/Participants`.

unaware of the state and health of the hardware it is running on cannot make intelligent decisions about its trustworthiness or appropriate use. This further suggests a need to address the *Granularity of Adaptation and Repair*.

- *Incomplete Information on Reliability and Weaknesses:* Current design methodologies make it difficult to understand the real sources of weaknesses in a design, leading to over-design of components that are not critical to reliability.
- *Analog and Passive Elements:* Analog, discrete, and passive components are critical to the data input and output paths on mixed-signal systems and can often be the weak link for data and control integrity. Systems that are unaware of the health of these components and unable to compensate for their failure or changes will have limited reliability.

The common causes of these challenges suggest a pressing need to develop a better **scientific and engineering** understanding of *information sharing and exploitation across the multi-layer system stack that supports computations*. Specific components of this inquiry should include the following set of fundamental questions:

1. How do we design hardware and software organizations that are prepared for repair?
2. What is the right amount of error filtering at each level in the stack—from circuits through application software—and what are the best techniques for filtering?
3. How do we formulate, analyze, and manage multilevel trade-offs for fault mitigation that generalize the idea of hardware-software trade-offs, including the interfaces for cross-layer information sharing?
4. What is a theory and design pattern set for efficient, light-weight error checking that exploits the high-level properties of hardware architectures, software applications, and algorithms?
5. What is a theory and practical framework for expressing and reasoning about differential reliability, including both application needs and hardware/system organization to meet those needs?
6. How do we design scalable system solutions that can adapt to varying error rates and reliability demands?
7. How do we design components and systems that degrade gracefully and systems that are aware of their overall health?

Developing this understanding will directly impact many key **national missions**, including:

- **Supercomputing**: In order to build the ExaFLOP supercomputers that will drive science, defense, and commerce in the 2020's while staying within the power budgets of major data centers, we must increase computational efficiency more than 250-fold while simultaneously drastically reducing the fraction of time spent handling errors and extending mean-time-between-failures (MTBF) to months or years. Architecture redesign and software cooperation will be necessary to achieve these goals.
- **Satellites**: Satellite technology supports cable television, in-theater warfighters, and science. On-board processing is necessary to optimize the limited communication bandwidth to the ground, but must operate within power limits of 20–30 Watts. As radiation-hardened technologies lag commercial technologies by several generations, we must find ways to use more advanced and energy-efficient commodity technologies without sacrificing system reliability.
- **Medical**: Reliable, ultra-low-power computing systems will enable a host of breakthroughs in medical technology, including personal genomics, sensing systems that help compensate for blindness, assistive technologies for the elderly, and implantable devices that operate for years without failure or the need for recharging.

- **Commercial Industry** Our commercial and financial infrastructures all depend critically on reliable computation. The current economic recovery demands that reliable computation continue in spite of financial austerity measures.
- **Transportation**: Advanced safety features and drive-by-wire control demand the greater computational capabilities of advanced technologies, but also have even higher safety requirements to safeguard human life.
- **Security**: Cross-layer reliability techniques provide the foundation necessary to support the security needs of electronic commerce, electronic medical records, and military applications.

A **research program** in the nascent area of cross-layer reliable systems could have tremendous impact and influence over the development of this critical technology. Because of the cross-cutting nature of this area, it will be essential for this program to enable collaboration of researchers with a wide range of expertise. The program should focus on developing example systems and the standard, open platforms that will enable the subsequent engagement of a larger community of domain experts across academia, government, and industry. It should also support and encourage the development of tools to model and characterize cross-layer reliable systems.

**Government leadership is essential.** The work necessary to achieve cross-layer reliable systems crosses the entire computing system ecosystem from integrated circuits to software applications. Therefore, no one vendor or research laboratory will be able to effect change by themselves. Wide-scale cooperation across specialties and organizations is necessary to revolutionize computing systems in this manner, otherwise the community will be facing yet another stop-gap revision that will only postpone these problems for a few more years! United States defense and civilian infrastructures—communication, finance, transportation, health—all depend critically on reliable operation, and the *government plays a key role in funding and providing pieces of this infrastructure, and in advocating and enforcing standards to enhance consumer safety.* Computer Engineering curricula must change to equip young engineers and programmers to design and develop reliable computing systems. Furthermore, Moore's Law scaling and the creation of value through new capabilities harnessing advanced computing have been a key **engine of economic growth** raising our standard of living in the United States. The research outlined above will provide US companies with the technologies required to sustain innovation and develop reliable products that will continue to bolster our economy and create high-value domestic jobs.

**Learn More** `http://www.relxlayer.org`

**Acknowledgment** This material is based upon work supported by the National Science Foundation (NSF) under Grant 0637190 to the Computing Research Association (CRA) for the Computing Community Consortium (CCC). The opinions, findings, conclusions, and recommendations expressed in this material are those of the study leaders and do not necessarily reflect the views of the NSF, CRA, CCC, or the employers of the study leadership or participants.

Contact E-Mail: `relxlayer@relxlayer.org` Document release number: LA-UR 10-08387