Model-based Computing: Harnessing the Power of Tera-scale Systems

The Future of Computing is Model-Based

The performance of future tera-scale computers, with tens or hundreds of integrated processor cores, will empower tomorrow's technology with more human-like capabilities. Computers will be able to "think" in terms of models—digital "concepts" of people, places, and things found within streams of bits and bytes—enabling them to do far more sophisticated work for users. By building and refining models of what a chair looks like, what data makes up a financial portfolio, or how an object in a video game should move, future computers will be able to perform a range of complicated tasks, from organizing digital photo collections to testing new investment strategies to creating virtual characters for online worlds.

> **By using models, future computers will be able to perform a broad range of sophisticated tasks**, from organizing digital photo collections to testing new investment strategies to creating virtual characters for online worlds.

Computers will build models in a way that's similar to how humans form concepts. For instance, a child develops the concept of "chair" by observing many objects over time, and learning which are considered chairs and which are not. Once she has formed a concept, the child can recognize *any* chair and can synthesize what she's learned to create entirely new concepts of chairs. Similarly, a computer, fed sufficient data or examples, can "learn" and refine a model over time, and synthesize new models from what it has learned.

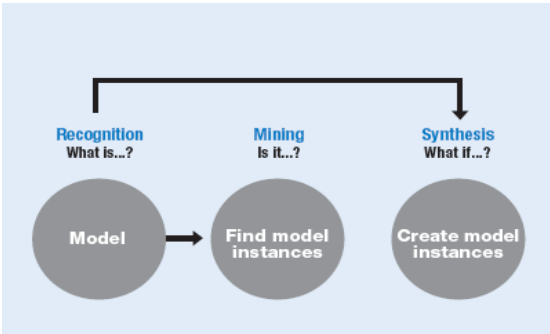
We will see the emergence of many model-based computing applications in the coming years. This emerging model of computing already has penetrated game consoles and some segments of servers, such as HPC (high performance computing) servers.

Intel researchers are exploring the demands of these powerful new applications in order to guide the design of future tera-scale computer architectures.

Model-based Computing Techniques: Recognition, Mining and Synthesis

Angela's parents are about to celebrate their 25th wedding anniversary. Using a model-based application, Angela searches through thousands of images and videos stored on the hard drive of her desktop PC. Within minutes, the application finds more than 50 photos and video clips that include the couple. Angela will sort through these and present her favorites at the anniversary celebration. She also decides to use her software and the data it has mined to create an image of what her parents might look like on their 50th anniversary, and to alter a video of the couple walking down the aisle on their wedding day, to show them as they appear today.

To enable applications such as the one described above, computers must be able to develop, manipulate and evaluate mathematical models. These capabilities can be subdivided into three categories: recognition, mining and synthesis (RMS). ***Recognition*** involves identifying a set of data within bits and bytes that constitutes a model, then building the model. ***Mining*** refers to searching a database, such as the Web or a photo collection, to find data that matches the model. ***Synthesis*** involves creating a new model by combining model data in new ways.



Recognition, Mining and Synthesis (RMS) techniques can be used to create a model, find instances of that model, and create or "synthesize" new models.

Model-based applications go through an iterative process of refining models, feeding the results of mining and synthesis back to the building process (recognition). Today human interaction is needed to guide the refinement process, but the ultimate goal is to remove humans from the process to the extent possible, in order to accelerate results. This could be especially useful for some model-based applications. For instance, an application that could refine investment models on its own could make millisecond stock trades possible, as no (relatively slow) human keystrokes would be involved. A difference of a millisecond could translate to millions of dollars for some financial transactions.

Emerging Model-based Applications

Model-based computing techniques and tera-scale capabilities will enable a broad range of powerful new applications for desktop and notebook computers. Following are just a few of the possibilities.

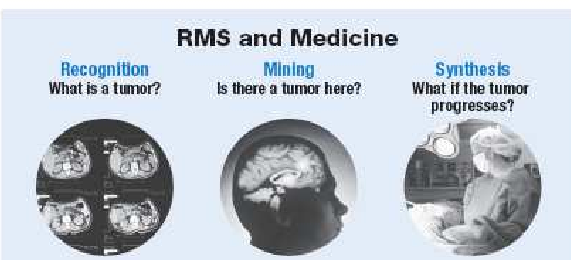
Entertainment Simulations Model-based applications could enable a far richer Web experience. As consumers gain access to libraries of music, film and television shows, model-based applications could unleash new channels of creativity. Users could put their friends or family members into a movie, or model to see how a different pitcher might have changed the outcome of a pivotal baseball game. Film directors could use MBC techniques to test how different actors might appear in a future movie.

Virtual Worlds Virtual reality will reap the benefits of RMS techniques and tera-scale capabilities. This combination will bring far greater realism to applications such as Second Life, the online 3D virtual world that more than seven million "residents" worldwide create and own. For instance, RMS techniques could enable computers to recognize human movement, enabling users to control their virtual personas without a keyboard. And techniques such as ray tracing, used to model lighting effects such as reflections and shadows, will bring much greater detail to virtual scenes, and enable the creation of more realistic avatars. Such advanced modeling capabilities will make the virtual worlds of the future far more engaging and immersive.

RMS techniques, combined with tera-scale computing capabilities, could enhance the realism of virtual scenes such as this one, from Second Life.

Personal Investing In selecting equities, investors should analyze a broad spectrum of factors that could include not just company financials but industry trends, oil prices, currency rates, and more. There is so much data to consider that in the end, investors often rely simply on a stock's price-to-earnings ratio or recommendations by friends or brokers.

Model-based applications could radically change that. Using a model of a successful equity investment, individual investors could mine a data set to identify other potential successful investments and create "what if" scenarios that consider various investment periods and the potential effects of such events as a rise in interest rates or a corporate acquisition. It's impossible to take all the uncertainty out of an investment of course, but a model-based application could give investors much more information to help with their decisions.



A model-based application could be used to detect tumors. *Images courtesy of Surgical Planning Laboratory, Department of Radiology, Brigham and Women's Hospital.*

Improving Healthcare A model-based medical application could build a model of a tumor by mining patient data. The application could then synthesize the data to predict the effects of the tumor's progression for a particular patient, and whether treatment is advisable. The application also could be used to determine the efficacy of various treatment options, analyzing what is known about each treatment model in conjunction with the patient's medical history and condition. For instance, the application could simulate the effects and prognosis of treating a cancer tumor using chemotherapy, radiation, surgery, a combination of approaches, or no intervention at all. By examining the outcomes of each treatment, the doctor and patient could decide on the best course of action.

Surveillance Today's surveillance cameras simply record what they see. Imagine if those images were run through a computer that could recognize security risks and alert authorities. For example, a bank's surveillance system, trained through models to recognize a gun or suspicious behavior, could immediately alert the bank or the police when someone in the bank is carrying a gun or acting suspiciously.

On a larger scale, a model-based application could help to monitor cargo shipments. Today, only a small percentage of container traffic on ships is inspected. Using data from RFID tags, surveillance cameras, biological sensors and other monitoring systems, a computer could monitor cargo using a database of models to identify which shipments require an alert and further inspection.

System Design for Model-based Computing: The Power of Convergence

Virtually all model-based software applications, from fields as diverse as finance, medicine and consumer entertainment, map to a very small, common set of computational algorithms. These include techniques such as collision detection, which is used to determine if objects in a virtual world are colliding, and Monte Carlo methods—sophisticated algorithms that predict the behavior of systems, and which are widely used in financial applications for tasks such as pricing complex options derivatives.

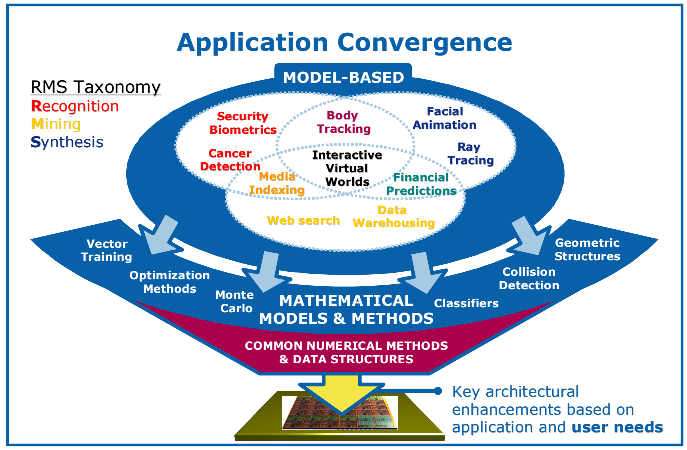
Because of this convergence of computing techniques, systems designers can support a broad range of model-based applications by tackling a small set of design challenges.

Convergence also means that future Intel architectures can be optimized to run these algorithms most efficiently, optimizing application performance. In addition, as a result of convergence it may be possible to cost-effectively design a single architecture that will meet the needs of most or all model-based applications.

Design Challenges

The requirements of many model-based applications will approach or exceed one teraflop—one trillion floating point operations per second. Such applications can benefit from the orders of magnitude performance increases that tera-scale computers of the future will deliver. However, in order to achieve high performance from these systems, several hardware and software challenges must be overcome.

For instance, tera-scale computers will rely on multiple "threads"—separate streams of operations performed in parallel by different processors. Such parallelism accelerates results, but it can be difficult to manage communication and synchronization between threads. In addition, in model-based computing applications, the data to be processed often is scattered in a variety of locations. "Gathering" the data items would enable them to be processed using many fewer instructions, maximizing efficiency. Intel scientists are exploring hardware and software solutions to both of these challenges and others, to ensure that tera-scale systems will deliver on their promise of high performance.



Model-based applications share common software algorthims that could be optimized in future achitectures

Conclusion

Model-based computing will transform the way that consumers use technology in the future, bringing powerful new applications to future computers. Adding the appropriate capabilities to processors will help enable these new types of applications to be written and adopted for widespread use. Model-based computing techniques will enable real-time, detailed physical simulations, making games and virtual worlds much more accurate and interactive than they are today. Combining these techniques with tera-scale computing performance could enable entirely new kinds of applications that we have yet to envision.

Model-based applications share a small, common set of mathematical models and computational algorithms. This convergence has strong implications for future computer system developers. It means that a diversified set of emerging model-based applications, from market segments such graphics, gaming, media mining, financial analytics, and interactive virtual communities, presents a relatively small set of common platform challenges. A general-purpose processing platform designed to address these challenges has the potential to significantly increase programmer productivity and enhance the user experience.