

SUPER SOC: PUTTING THE WHOLE (AUTONOMOUS) SYSTEM ON THE CHIP (ASOC)

Michael J. Flynn

Gates Computer Science Bldg., Rm. 335
353 Serra Mall
Stanford, CA 94305-9030
Email: flynn@ee.stanford.edu

ABSTRACT

With dramatic advances in transistor density, it's time to look ahead to the completely autonomous system on a single die (ASOC). This represents a convergence of RFID type technology with SOC silicon technology coupled with silicon transducers, sensor controllers and battery, all on the same die. The major architectural implication is design for extremely low power (1 microwatt or less) and strict energy budget. This requires a rethinking of clocking, memory organization, and processor organization. The use of deposited thin film batteries, extremely efficient RF, digital sensors and MEMS complete the ASOC plan.

1. INTRODUCTION

While system on a chip (SOC) technology represents an expanding part of the microprocessor marketplace; growing at 20% per annum rate, there's much more to come [1].

The typical SOC consists of multiple heterogeneous processors and controllers and several types of memory (ROM, cache, eDRAM, etc.). The various processors are oriented towards one or more types of media processing. Typical applications include cell phones, digital cameras, MP3 players and various gaming devices.

Another fast growing chip marketplace is autonomous chips. These have little processing power or memory but have RF communications and some type of self contained power source or power management. The more elaborate autonomous chips also include (or are coupled with) some type of sensor. The simple versions include RFID chips [2], smart cards and chip implanted credit cards. The simplest is the passively power RFID. The chip simply reflects the source RF carrier and modulates it (using carrier power) to indicate its ID. More complex examples include patient monitoring alarm [3] and the Smart Dust DARPA research program [4, 5] of the '90s. Both of these used battery powered RF to broadcast an ID on a detected sensor input.

<i>System</i>	<i>Passive id</i>	<i>Active id</i>	<i>RF sensor</i>	<i>ASOC</i>
Example	RFID, Smart Card (simple)	Smart card, Active RFID	Smart Dust; RFID + sensor	
Power source	None	Short term battery	Battery	Integrated Battery
Maximum Memory	ROM ID (1KB)	R/W ID + parameters (2 KB)		R/W extensive (100 MB)
RF range (meters)	Passive; order of cm	Active 1-10	10 - 20	10 +
Compute	None	FSM	FSM	1 or more CPU

Table 1. Some autonomous and ASOC examples (FSM represents a simple finite state machine or micro controller).

The various Smart Cards and Money Cards include VISA cards and Hong Kong's Octopus Card. All (except those that require contact) use a form of RFID. The simplest cards are passive without on-card writeable memory.

Written in the fond remembrance of our dear colleague, Stamatis Vassiliadis. I also acknowledge the help of Professor Wayne Luk and Dr. Patrick Hung in preparing this paper.

Records are updated centrally. Implementation is frequently based on Java Card [JAV]. Based on the extraordinary interest there are a series of contactless identification cards (RFID) standards:

- ISO 10536 close coupling cards (0 - 1 cm)
- ISO 14443 proximity coupling cards (0 - 10 cm)
- ISO 15693 vicinity coupling (0 - 1 m)

The future autonomous SOC or ASOC is the combination of the SOC with the AC (autonomous chips) technology (Table 1). While conceptually simple the engineering details are formidable as it involves rethinking the whole of processor architecture and implementation to optimize designs for low and sub microwatt operation.

The motivation for ASOC follows the Smart Dust [4] which started in the early '90s and pioneered significant work in the sensor and RF areas. That project targeted sensor plus RF integrated into a form factor of the order of 1 mm³ called motes. As a power source it relied on AA type batteries. That project was sponsored by DARPA and was target at sensing an "event"; a moving object, a thermal signal, etc.

ASOC is an updated extension of that work that places more emphasis on computational ability and memory capacity; as well as fully integrating a power source on die.

A simplified classification of AC (autonomous chips) is:

1. Simple identification of the die itself (as in RFID) with RF response.
2. Identification of a sensor detected "event" with RF response (as in Smart Dust and many Smart Cards).
3. Detection of an "event" and processing (classification, recognition, analyzing) the event (ASOC) with RF response of the result.

The ability of the ASOC to process data is clearly valuable in reducing the amount of sensor data required to be transmitted. It enables applications (such as supporting planetary exploration) where interactive computational support is impossible; so too with recognition of rare bird or other species in remote areas; or swallowing an ASOC "pill" for diagnosis of the gastro intestinal tract. Not all dimensions of ASOC are equally important in all application. A rare species "listening" post may require little size concern and may have ample battery support. We look at ASOC as a toolkit for the new systems designer; offering the ability to configure systems to respond to an almost endless set of environmental and computational concerns.

In this paper we consider the evolution of silicon technology, limits on batteries and energy, architecture implications, communications, sensors and applications.

2. TECHNOLOGY

The whole revolution in silicon chip technology has been powered by the ever increasing transistor density. Over the next 10 years transistor and memory density is expected to increase 10 fold [1] to more than one billion transistors / cm². Since a reasonable powerful processor can be realized with a few 100,000 transistors, there are a lot of possibilities for ASOC applications.

The main problem for useful ASOC is battery power or stored energy. In dealing with this issue recall two general relationships, relating silicon area, A , algorithmic execution time, T , and power consumption, P (in these expressions k is a constant):

$$AT^2 = k \tag{1}$$

This well known result [6] simply related area (the number of transistors) to the execution time required to complete an operation. The more area (transistors) used the faster (smaller) the execution time.

$$P^3T = k \tag{2}$$

This result [7] is based on voltage scaling arguments. It's easy to see that as voltage is decreased power is reduced by the square but speed is reduced linearly. But in transistors the charging current (representing delay) and voltage have a non linear relationship. This gives the cubic result. We can rearrange this as:

$$F_2/F_1 = (P_2/P_1)^3$$

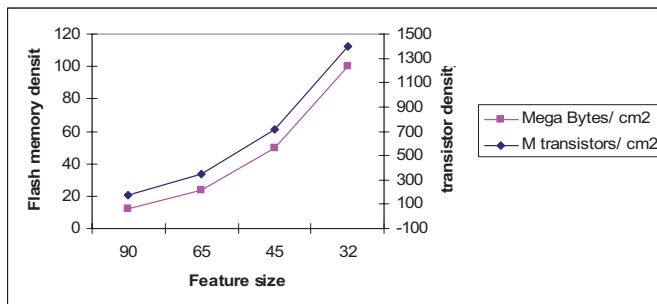


Fig. 1. ITRS [1] projection for transistor density.

So if we want to double the frequency we should expect the design to use 8 times more power. While the range of applicability of expression 2 is not precise, suppose we use it to project the frequency of a processor design that operates at a microwatt. The best power-performance design of today might consume one watt and achieve one Gigahertz (corresponding perhaps to 1,000 MIPS); this may be optimistic. Reducing the power by a factor of 106 should reduce frequency by a factor of 100 or 10 Megahertz. Within the past two years a sensor processor has been built that achieves almost 0.5 MIPS per microwatt [8]. While this is an order of magnitude away from our target of 10 Megahertz per microwatt, silicon scaling projections should compensate for the difference.

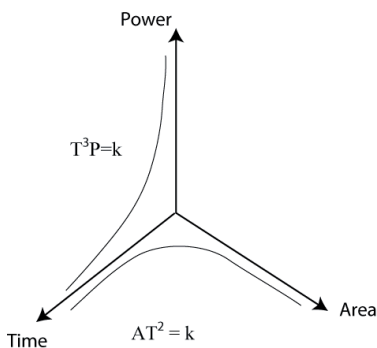


Fig. 2. The area - time - power tradeoff

3. POWERING THE ASOC

The key problems in forming robust ASOC are energy and lifetime. Both relate to the power source, i.e. the battery. Batteries can be charged once or rechargeable (with varying recharge cycles). For ASOC purposes, rechargeable batteries use scavenged energy from the environment. The capacity of the battery is usually measured in milliamp-hours; which we convert to Joules (watt-seconds) at 1.5 volt. Both capacity and rechargability depend on size which we assume is generally consistent with the size and weight of the ASOC die (about 1 cm² surface area).

In Table 2 we list 3 common battery types; the printed [9, 10] and thin film batteries [11] can be directly integrated into the ASOC die (usually the reverse side); button batteries are external and are less than 1 cm in diameter.

Printed batteries are formed by printing with special inks in a flat surface; thin film batteries are deposited on silicon much as the system die itself.

Energy may be scavenged from many sources (some are illustrated in Table 3); usually the larger the battery format the more the charge. Much depends on the system environment as to which, if any, scavenging is suitable.

Assuming ASOC consumption of 1 uw (when active); the operational lifetime between charges is plotted in

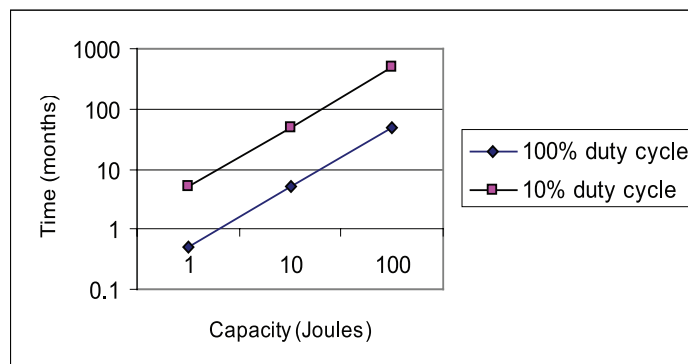
Type	Energy (J)	Recharge Y/N	Thickness (micron)
Printed	2 / cm ²	N	20
Thin film	10 /cm ²	Y	100
Button	200	Y	500 stand alone

Table 2. Batteries of ASOC.

Source	Charge rate	Comment
Solar	65 (milliwatts)/ cm ²	
Ambient light	2 (milliwatts)/ cm ²	
Strain and acoustic	A force (sound) changes alignment of crystal structure, creating voltage	Piezoelectric effect
RF	An electric field of 10V/m yields 16uW/cm ² of antenna	See [12]
Temperature difference (Peltier effect)	40 (microwatts/5°C difference)	Needs temperature differential.

Table 3. Some energy scavenging sources [13, 14, 15].

Figure 3. Duty cycle can play an important role in expending the ASOC serviceability. The assumption is that a passive sensor can detect an event and power up the system for analysis.

**Fig. 3.** Maximum time between recharge for 1microwatt of continuous power consumption.

Comparing Figures 2 and 3, if we can configure the ASOC to use of the order of 1 microwatt we should be able to incorporate a suitable battery technology especially if we have the ability to scavenge some addition energy.

4. THE SHAPE OF THE ASOC

The logical pieces of the ASOC die are shown in Figure 4. It consists of the power source, sensors(s), main computer and memory and the communications module. What distinguishes the ASOC from the earlier RFID + sensor technology is the compute power and memory. It is this facility that enables the system to analyze and distinguish patterns; to synthesize responses before communicating with the external environment.

Physically the ASOC is just a silicon die; probably 1 cm² in surface area. Surface size is dictated by cost which is determined by defect density. Current technology gives excellent yields for 1 cm² and smaller die sizes. Much below 1 cm² costs are limited by testing and handling so this represents the preferred size for most applications. Die thickness is limited by wafer fabrication considerations and is about 600 microns. A thin film battery deposited in the reverse side might add another 50 microns. The resultant ASOC would be 65 mm³ and weigh about 0.2

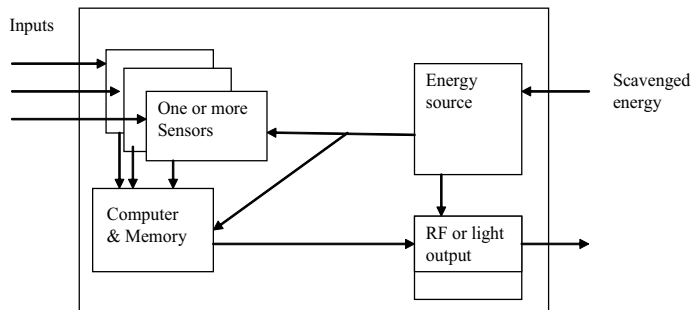


Fig. 4. An ASOC die.

grams. From Figure 1 it could have of the order of 1 billion transistors. These transistors would realize the sensors, computer, memory and RF; the battery is on the reverse side.

5. COMPUTER MODULE AND MEMORY

With a power budget of only 1 microwatt the micro architecture of the computer is considerably different from the conventional processor:

1. Asynchronous clocking: Data state transitions must be minimized to reduce dynamic power. There may be only one tenth asynchronous transitions required compared to a clocked system.
2. Use of VLIW: Transistors are plentiful but power is scarce, so the effort is to use any available parallelism to recover performance.
3. Beckett [16] has shown that by careful device design (lowering drive current and managing leakage) it is possible to arrange for overall power to be a reducing function of area. This sacrifices maximum operating frequency but the additional area can more than compensate by parallelizing aspects of the architecture.
4. Minimum and simple cache system: The memory and processor are in relatively closer time proximity. The processor is performing an action once every 0.1 microsecond; the FLASH memory has access time of 1 - 10 microseconds. A small Instruction cache and explicitly managed data buffers seem most suitable in the context of specified applications.

The Flash memory is another essential piece of the system as it has a persistent data image even without power. Current densities (NAND based Flash) give excellent access times, 10 microseconds, and ASOC capacity of perhaps 16- 64 MB.

As the technology is currently configured Flash is largely incompatible with integrated CMOS technology and seems restricted to off die implementations. However there are a number of Flash variants that are specifically designed to be compatible with ordinary SOC technology. SONOS [17] is a non-volatile example, and Z-RAM [18] is a DRAM replacement example. Neither seems to suffer from the conventional Flash rewrite cycle limitations (the order of 100,000 writes).

Even though the Flash memory consumes no power when it's not being used, when it is accessed the power consumption is proportional to the active memory array size[BEN]; i.e. the number of memory cells connected to each bit- and word-line (assuming 2-D square structure). In the context of ASOC this implies a memory partitioned into smaller units which may be most effective from both a power and access time basis.

6. RF OR LIGHT COMMUNICATIONS

One of the great challenges of ASOC is communications. There are two obvious approaches: communications laser and RF.

Lasers

Integrated laser with silicon is an emerging technology. A recent development [19] uses an Indium Phosphide laser with silicon waveguide bonded directly to a silicon chip. Using lasers for optical free space communications has possibilities and difficulties.

Optical sensors are quite responsive (figure from [20]); reception of 1 microwatt supports about 100 MHz data rates. The difficulty is that reception is subject to ambient light (noise). In general the signal must be 10 times greater than the noise. The other difficulty is beam divergence (especially in laser diodes). This requires optics to collimate the beam for low divergence [21].

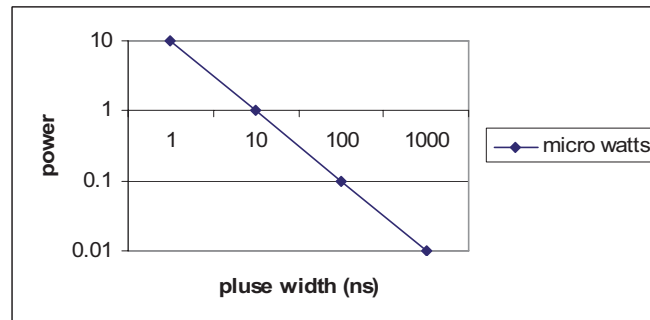


Fig. 5. Photo detector sensitivity is a function of pulse width.

Given the limitations, the use of laser free space (as distinct from fiber optics) for communication is probably a secondary prospect for ASOC.

RF

The work of the Smart Dust program seems to be the most relevant and useful here [4, 5]. That program demonstrated the integration of low power RF into an SOC chip. To summarize some of their many findings:

1. A feasibility study realized a transceiver achieving 100 Kbps over a 20 meter distance with an energy budget of 25 nJoules/bit. This corresponds to about 10^{11} bits / Joule/ Meter. One Joule of battery energy allows 100 Gbits to be transferred across one meter [5].
2. Communications with less than 1 milliwatt was not only feasible but likely to be commercialized. With typical duty cycle of less than 1% the average power consumption was between 1 and 10 microwatts.
3. There is a large data packet overhead (including startup and synchronization, start symbol, address, packet length, encryption and error correction). Short messages can have as little as 3% payload packet efficiency. It is better to create fewer longer messages.
4. As a result of (2) and (3) the system designer will want to minimize the number of transmission and maximize the data packet payload.

Sensors (vision)

Vision and motion sensors are usually configured as an array of photodiodes, with array sizes varying from 64 x 64 to 4k x 4k or more [22]. Each photodiode represents a pixel in the image (for grayscale digital images). Three or more diodes are needed for colored and multiple spectrum images. To conserve power and reduce state transitions an ASOC would probably implement the vision processor as an array with a single element per pixel.

In either image recognition or motion detection and recognition it's necessary to find the either the correspondence with a reference image or the direction in which a block in one image shifts with respect to a previous image. Determining the match between two images or successive frames of the same scene requires that the image be partitioned into blocks. The blocks of one image are compared to the reference or previous image block by

block in a spiral pattern. Each comparison involves computing the SAD index (sum of absolute difference). When the image configuration with the minimum SAD index is found the recognition or motion flow is resolved. While image recognition should be possible in milliseconds, the challenge for vision sensors is to meet the computational requirements of relatively fast moving objects.

While it's clear that the image sensors can be integrated in the ASOC, optics for focusing distant object in varying amounts of light can improve performance.

Sensors (audio)

As mentioned above the piezoelectric effect applied to silicon crystal can be used to record sounds and is the basis for many simple sound detection and microphone systems. Alternatively in specialized applications such as hearing aids it is sometime important to mimic the action of the ear. Various cochlear chips have been realized using a sequence of low pass filters to emulate the operation of the cochlea. In one silicon implementation [23] 360 cells each containing dual low pass filter were arranged as a linear array. Cochlea type implementations are usually preferred when speech recognition is required.

Since audible sound frequencies are relatively low there are few real time constraints for an ASOC.

Motion and flight

Of course the ultimate ASOC can both move and fly. Given a weight of only 0.2 grams motion per se is not a problem when the ASOC has associated mems. Memes and nano motors are used to anchor and move the ASOC across a surface. The energy required to move on a surface is simply the force to start (accelerate) and then to overcome friction. One Joule of energy translates into 10^7 Ergs. An erg is the energy required to move gram for 1 cm with the force of a dyne. So slow motion (order of 1-2 cm/second) that occurs relatively infrequently (less than 1% duty cycle) should not cause significant ASOC energy dissipation.

The motion of flight is by far the most complex. Various attempts [24] have been made for small vehicle autonomous flight. Flight encapsulates many of the ASOC challenges: power, vision (avoiding obstructions), environment (wind, etc) and communications. While the flying ASOC is a long way off; such systems are feasible as any small fruit fly [25] knows!

7. CONCLUSIONS

There's a whole new field to be explored based on the next generation of autonomously based system on a chip (ASOC). As we have seen, transistor density improvements will enable close to a billion transistors per cm^2 . This enormous computational potential has a major limitation: limited electrical energy. There is a new direction opening in computer architecture, *nano computing*, to contrast with historical efforts in supercomputing. The target of this field is to produce the algorithms and architectural approaches for high performance at less than one millionth current levels of power dissipation; freeing the chip from external power coupling.

For untethered operation a form of wireless communication is required. This is another significant challenge especially with a power budget also in the order of microwatts. While RF is the conventional approach some form of light (or infrared) may offer an alternative.

Finally, digitizing the sensors and even the transducers offers a final challenge where multiple sensors are integrated into a seamless SOC.

The best designs anticipate system complexity and deal effectively with the unanticipated. And system complexity includes many issues overlooked in this short paper: component design and suppliers, design tools, validation and testing, security, etc. Successful tradeoffs across a myriad of issues define effective design.

While there's little expectation that all of the ASOC components discussed here will actually be integrated into a single die, there are many different combinations. Each combination with its own system requirements must be optimized across all of the constituent components. The designer is now no longer concerned about a component but only about the final system and becomes the ultimate *systems engineer*.

8. REFERENCES

- [1] ITRS roadmap summary 2006, <http://www.itrs.net/Links/2006Update/2006UpdateFinal.htm>.
- [2] RFID, RFID Journal, <http://www.rfidjournal.com/>.
- [3] Patient monitoring University Hospital of Ghent, <http://www.rfidjournal.com/article/articleview/3120/1/1/>.
- [4] B. W. Cook, S. Lanzisera, and K. S. J. Pister, "SoC Issues for RF Smart Dust," *Proc. Of the IEEE*, vol. Vol. 94, No. 6, June 2006.
- [5] A. Molnar et al, "An ultra low power 900 mhz rf transceiver for wireless sensor output," in *Proc. Custom Integrated Circuits Conference*, 2004, pp. 401–404.
- [6] J. D. Ullman, *Computational Aspects of VLSI*. Computer Science Press, 1984.
- [7] M. J. Flynn, P. Hung, and K. W. Rudd, "Deep-submicron microprocessor design issues," *IEEE Micro Magazine*, pp. 11–22, July-August Issue, 1999.
- [8] B. Zhai, L. Nazhandali, J. Olson, A. Reeves, M. Minuth, R. Helfand, S. Pant, D. Blaauw, and T. Austin, "A 2.60pJ/Inst Subthreshold Sensor Processor for Optimal Energy Efficiency," in *IEEE Symposium on VLSI Circuits, 2006. Digest of Technical Papers*, 2006, pp. 154–155.
- [9] PowerID, Power Paper Corp., <http://www.powerpaper.com>.
- [10] V. L. Pushparaj et al, "Flexible Nanocomposite Thin Film Energy Storage Devices," <http://www.newswise.com/articles/view/531241/>.
- [11] Cymbet Corp, The POWER FAB (Thin Film Lithium Ion Cell) battery system, <http://www.cymbet.com>.
- [12] E. M. Yeatman, "Advances in power sources for wireless sensor nodes," in *Proceedings of 1st International Workshop on Body Sensor Networks*, London, 2004.
- [13] S. Meninger et al, "Vibration-to-electric energy conversion," *IEEE Trans. VLSI Systems*, vol. Vol. 9, No. 1, pp. 64–76, Feb. 2001.
- [14] S. Roundy et al, "Power sources for wireless sensor networks," in *Proc. 1st European Workshop Wireless Sensor Networks*, Jan. 2004, pp. 1 – 17.
- [15] L. D. Partain, *Solar Cells and Their Applications*, ser. Wiley Series in Microwave and Optical Engineering. J. Wiley, 2004.
- [16] P. Beckett and S. Goldstein, "Why area might reduce power in nanoscale cmos," in *IEEE International Symposium on Circuits and Systems 2005 (ISCAS 2005)*, vol. 3, Kobe, Japan, May 2005, pp. 2329–2332.
- [17] T. Sugizaki, M. Kobayashi, M. Ishidao, H. Minakata, M. Yamaguchi, Y. Tamura, Y. Sugiyama, T. Nakanishi, and H. Tanaka, "Novel multi-bit sonos type flash memory using a high-k charge trapping layer," in *IEEE Symposium on VLSI Technology, 2003. Digest of Technical Papers*, 10-12 June 2003, pp. 27–28.
- [18] Z-RAM Innovative Silicon Corp., <http://www.innovativesilicon.com/>.
- [19] A. Fang, R. Jones, H. Park, O. Cohen, O. Rada, M. Paniccia, and J. Bowers, "Integrated hybrid silicon evanescent racetrack laser and photodetector," in *12th OptoElectronics and Communications Conference*, Yokohama Kanagawa, Japan, Jul. 2007.
- [20] D. Johnson, "Handbook of Optical through the Air Communications, www.imagineeringezine.com."
- [21] R. Paschotta, "Encyclopedia of laser physics and technology, www.rp-photonics.com/encyclopedia.html."
- [22] T. Komuro, S. Kagami, and M. Ishikawa, "A dynamically reconfigurable simd processor for a vision chip," *IEEE Journal of Solid-State Circuits*, vol. Vol. 39, No. 1, pp. 265–268, January 2004.
- [23] B. Wen and K. Boahen, "Active bidirectional coupling in a cochlear chip." in *Advances in Neural Information Processing Systems*, 2005.
- [24] J. Zufferey and D. Floreano, "Toward 30-gram Autonomous Indoor Aircraft: Vision-based Obstacle Avoidance and Altitude Control," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA 2005)*, 18-22 April 2005.
- [25] S. Sane, "The aerodynamics of insect flight," *The Journal of Experimental Biology*, vol. 206, pp. 4191–4208.
- [26] Java Card Platform Specification, <http://java.sun.com/products/javacard/specs.html>.
- [27] L. Benini, A. Macii, and M. Poncino, "Energy-Aware Design of Embedded Memories: A Survey of Technologies, Architectures, and Optimization Techniques," *ACM Transactions on Embedded Computing Systems*, vol. Vol. 2, No. 1, pp. 5–32, February 2003.