"You have to discover and define your problems in order to solve them," says Bill Courtright. He’s explaining the rationale behind the new Data Center Observatory (DCO) at Carnegie Mellon. The DCO is a working data center — with space for 1,000 rack-mounted servers — which also provides a real-life testbed for observing, analyzing and attacking complex operating issues. The goal: to get a better handle on skyrocketing costs faced by data-center managers worldwide.

Courtright, executive director of the Parallel Data Lab (PDL), is an alumnus with a track record as an entrepreneur in the industry. The DCO was conceived and is co-led by Greg Ganger, an Electrical and Computer Engineering professor. Ganger says the idea grew from annual retreats with industry that he’d been organizing as head of Carnegie Mellon’s Parallel Data Lab, a hub for research on storage systems.

“People were saying, ‘storage management is killing us,’” Ganger recalls, but they couldn’t be much more specific: “The administrators just knew that they were busier than ever and the costs kept going up.” The hassles weren’t confined to managing the storage servers that hold databases and files. A data center also may have hundreds, or thousands, of commodity machines that are organized into “clusters” for computation tasks: large-scale computer modeling and simulations, for instance.

And a meta-issue with all of the machines is that operating costs far exceed the purchase costs. Thus while computers have gotten much cheaper and more powerful, the savings are now being eaten up.

The general reasons for this are well known. As higher-performance machines are packed more densely into racks, they consume more electricity, not only for power but for cooling (machine room air conditioning is a monstrous task). Over the typical service life of the machines, the electric bills alone now equal or surpass the purchase prices, and energy costs are rising. The human costs of dealing with increased complexity, meanwhile, run higher still. At many data centers, administrative expenses are now four to seven times the annual capital expenses for new equipment.

Zeroming in on problems has been difficult, however. The everyday demands on data centers make it hard to even track precisely what the staff’s time is being spent on, let alone engage in radical cost-cutting experiments. As Courtright puts it, data centers are critical operations “trying to do research on the side” — which hasn’t worked well.

**What is a ‘Data Center Observatory’?**

It’s a new kind of university research center, tackling the world’s data problems in a live-action setting.

**By Mike Vargo**

Over the service life of these machines, the electric bill to power and cool them will equal or surpass their purchase price.
This is what led Ganger to propose the DCO, a data center where the study of human factors and other issues could be built into the mission, “a place to measure everything and try new solutions.” The DCO opened this past summer and support on campus is strong. The computing work being processed at the DCO comes from Carnegie Mellon. It serves as an added resource for researchers across the university.

In the past, Ganger explains, many faculty and graduate students have had computing equipment paid for by research grants or donated by firms, “but the operating costs aren’t covered. They may wind up with a dozen machines they have to find a place for, and they have to become system administrators, or hire administrators, to get their research done. As we transition this computing to the DCO, we can ease the burdens on researchers and save Carnegie Mellon a lot of money.”

The Data Center Observatory is housed in the new Collaborative Innovation Center, perched above Panther Hollow along Forbes Avenue. Bill Courtright seems an ideal person to be involved, having a Ph.D. in electrical and computer engineering from Carnegie Mellon plus strong industry experience. In 1999, with ECE professor Garth Gibson, he co-founded the now-global storage systems firm Panasas. Courtright was chief operating officer before stepping down to take on this new challenge.

Strolling through the DCO in late summer, when the first rows of servers had been installed and were starting to be used, Courtright eagerly showed off its features. He noted the energy-efficient cooling system from APC (American Power Conversion, a lead industry partner), which places small cooling units in the rows among the servers instead of relying on chilled air blown up from floor vents. He described how the DCO was designed for controlled shutdown in case of a power outage, to avoid loss of work in progress.

Above all, Courtright waxes eager on the research to be done at DCO. Already the research team has begun to keep time logs that allow fine-grained breakouts of human costs: “We want to know, for instance, how much time goes into tending the computers vis-à-vis the networks — and then within networks, how much on installation, configuration, debugging?” Good information at these levels can vastly inform cost-control experiments, which will mix new engineering and management approaches.

Courtright says the tinkering may be disruptive at times, but planned to minimize interference with research computing. Indeed, one experiment (in conjunction with APC) will involve dynamic scheduling and balancing of jobs, so as to use the fewest possible servers at any time: those not in use could be powered down to save electricity.

Altogether the Data Center Observatory is a massive joint project of the College of Engineering, the School of Computer Science and CyLab. The DCO is also supported by government agencies and industry affiliates. Furthermore, early on-campus users like ECE professor Elias Towe’s research group in nanotechnology have begun to collaborate with the project. When fully built out the DCO will have about 1,000 servers in a 2,000-square-foot room, making it a moderate-sized but very densely loaded data center. It is of course not the only data center on campus, but is the only one of its kind thus far at any university, mixing real-time data processing with full-time observation and research. In short, it’s another key first for Carnegie Mellon.

“People were saying, ‘storage management is killing us,’” Ganger recalls, but they couldn’t be much more specific: “The administrators just knew that they were busier than ever and the costs kept going up.”
How and why the new CenSCIR will embed sensing in pipelines, electric grids and other vital infrastructure.

By Mike Vargo

We have auto brake pads that “sing” to alert us when the lining is wearing thin; we have smoke detectors and cell phones that beep when their batteries run low. Wouldn’t it be nice to have equally good early-warning systems for corrosion in oil pipelines — or for the times when conditions in the power grid are conspiring toward a rolling blackout?

There are researchers at Carnegie Mellon who believe such systems could save many billions of dollars and much disruption every year. They are now pooling their talents in CenSCIR, the new Center for Sensed Critical Infrastructure Research.

Present methods of maintaining infrastructure are “fraught with inconsistencies,” says Jim Garrett, head of Civil and Environmental Engineering. They are a mixture of “infrequently scheduled visual inspections,” computer-based tools that don’t provide decision-makers the support they need, and policies of “fixing it when it’s broke or really expensive to fix. We need to move to proactive, consistent asset management” of roads and bridges, water and power services and more, he says. Such asset management requires large amounts of collected data upon which decision support systems can then be built.

Technologies to enable this are now being developed, notes José Moura, professor of Electrical and Computer Engineering. “People are able to make inexpensive, autonomous sensors about the size of a quarter,” he says. They can be embedded throughout any infrastructure to sense for various early signs of trouble — the cracking sound of tiny fractures, chemical or electrical signs — and then wirelessly transmit what they detect.

“We also need systems to collect and process the data, to make decisions,” Moura adds. And since Carnegie Mellon is involved at the cutting edge in all such work, he asks, what better place to pull the capabilities together into cost-effective solutions that would actually be used?
Garrett and Moura are co-directors of CenSCIR. They had been pursuing separate but closely related work — Moura with his colleagues in ECE and elsewhere, Garrett through a previous entity within the Institute for Complex Engineered Systems (ICES), called the AIS (Advanced Infrastructure Systems) Lab — until Jeanne VanBriesen, a CivE professor who specializes in urban water quality, suggested they merge their streams.

To quickly review the moving and shaking that has transpired since then: the former ICES AIS Lab has been folded into CenSCIR. Leading researchers from many departments are coming aboard, including from Engineering and Public Policy. Funding is coming from NSF, NETL (the National Energy Technology Laboratory) and others. An executive director from industry has been hired, Matthew Sanfilippo, formerly with Michael Baker Corporation.

CenSCIR as a whole is under the formidable wing of ICES. And …

Soon to come: Sensor Andrew. “Andrew made Carnegie Mellon the most wired campus, and Wireless Andrew made us the most wire-lessed. This will make us the most sensed campus,” Moura says. It will essentially “replicate Wireless Andrew, but with sensors. A few hundred at first, in a year or so, then thousands in the second phase.”

Sensor Andrew, like its forebears, will be both a practical system and a testbed for new technologies (such as the Firefly sensor “platform” developed by ECE’s Raj Kumar and others). Things to be sensed could include: water quality. Flows and potential leaks in various utility lines. Security and safety in certain areas. Movements and locations of certain people (with permission, and with privacy concerns addressed). Soundness of structures, electric power status, and so forth. The range of future applications in the rest of the world should be evident.

Aside from the technical challenges, two kinds of psychological hurdles must be cleared to win wide acceptance of massively sensed infrastructure in the so-called real world. One hurdle stems from disappointment with early systems. Jim Garrett says, “When you talk to infrastructure managers, time and again you will hear: ‘I tried sensing five years ago. I spent more time and money keeping the stuff calibrated, and maintaining it, than it was worth.’ And that’s not surprising. Until recently, sensors were expensive and capricious and riddled with wires.”

Concludes Garrett: “This time we’ve got to work on economics, and on things like the middleware that supports the creation of decision support systems. Today you can buy almost any kind of [wireless] sensor imaginable, but people won’t just put them out there just to generate more data. It’s the systems issues that must also be addressed. That is where CenSCIR is focused.”

And according to José Moura, the other hurdle involves a paradigm shift. Using the power grid as an example, he says the traditional approach to monitoring is: “Let’s try to describe and model in minute detail what’s going on.” This tends to involve a relatively few high-precision sensors, tracking what are perceived to be key variables at key points, with computers running big simulations and forecasts to guide human decision-makers at local command centers.

“The trouble is, you cannot model this highly complex behavior with the sophistication you need,” Moura says. A better paradigm, he says, may be “let’s go with simpler, coarser models but lots of sensing. Cheap sensors all over the place, each one maybe not so reliable, but there are so many of them. Then just try to abstract global behavior from all this local news. You may detect things at an early stage that you would otherwise miss.”

CenSCIR-affiliated scientists already have been working on, and in some cases filing for patents on, technologies ranging from ingeniously simple and cheap MEMS-based sensors to middleware and modeling. With the toolkit and the minds now being assembled, the future can literally be sensed.