Data Space

CLOG

Your Activities Leave Many Footprints, But Are Just One Part Of What Is Leading Our World Into The New Generation Of Big Data

You're sitting in the passenger car of a train, it's rumbling down the tracks and you wonder if you'll make it to work on time. Grabbing your smart phone from your bag, you send a quick text to your boss—"train running behind, going to be late." But you figure you've got time to kill, so you take out your laptop and catch up on a couple of reports that are due. You make a quick post on Facebook, a comment on Twitter.

This vast amount of information is only continuing to grow, to accumulate, and challenge traditional data management approaches. It's leading us to search for new solutions to what we call the Big Data problem—or, in the case of IBM, the Big Data opportunity.

There are two types of Big Data: *Data at Rest* and *Data in Motion*. A large volume of data that isn't immediately analyzed as it comes in, is *Data at Rest*. IBM offers two approaches for handling this data: large scale data warehouses (like Netezza) for structured data, and BigInsights (IBM's Hadoop-based offering) for data with a variety of structures.

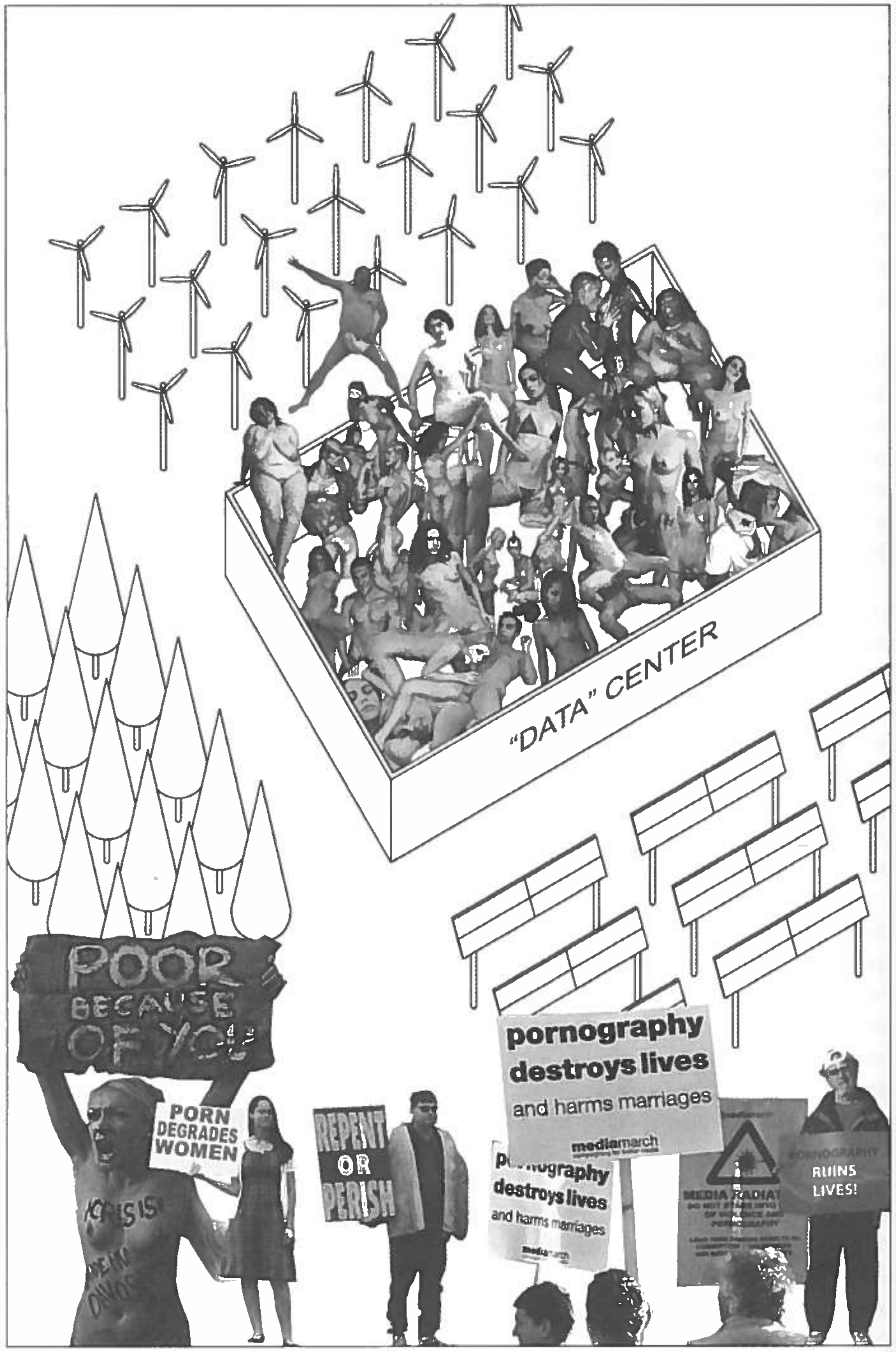
Data in Motion addresses the issue of analyzing data before it reaches the source. For example, sensors attached to a premature baby generate over a thousand unique data items every second, and by analyzing this data as its being generated, doctors can make diagnoses up to twenty-four hours sooner.

IBM wants to be able to encompass Big Data analytical techniques without ignoring or eliminating traditional solutions like data warehouses. The solution is the IBM Big Data platform with architecture to support hardware or application failure. It is also extremely scalable, and is able to run on many nodes to increase performance. One of the most important aspects of the Big Data Platform is integration. IBM's governance and integration products allow secure management a high flow of data exchange between wide varieties of sources. This enables IBM to combine the benefits of data warehouses and analytics, and conquer both new and traditional information sources.

Businesses will take advantage of insights they can gain from new sources of data, like the data being generated from the train, to make sure you arrive safely. They're analyzing it as it comes in using stream computing. Others, like telecommunication companies, are storing your cell phone logs. That text message (and its meta data) you just sent to your boss will be stored for a whole year—just one more bit of data that's being added to many terabytes. And others are currently crawling Twitter for your comments, your feeds, and they're analyzing it. What products are you most likely to buy? Do you favor certain brands over others? How can we best target your audience for marketing? These are just some of the many questions they're asking and through Big Data they can now get the answers. In the technological world, your activities leave many footprints, but are just one part of what is leading our world into the new generation of Big Data.

And this is just the beginning.

KAYLA MARTELL



While architects seem to be trying to make architecture more accessible, public, open, and glassy, somehow it seems we are simultaneously still keeping our global sex drive under the covers.

Larry Sultan, an influential Californian photographer spent a portion of his career photographing porn sets in Los Angeles's San Fernando Valley—what has been deemed the Silicon(e) Valley of the porn industry. His photographs are full of subtle narratives all centered around the middle class homes rented out for short periods of time for use in pornographic material. In this way, he captured the space where pornography is made. But now that a majority of pornography is distributed digitally, where is this data stored? Not the B-52's l-l-l-love shack, but big business data centers.

As of 2006, porn is a global business estimated at about $100 billion annually or over $3,000 per second. Every second over 28,000 users are viewing pornography and 372 users are typing adult search terms into search engines.[1] While these are the latest statistics, based on Google Trends, the amount of searches for terms such as "porn," "porno," "pornos," and "porn videos," have increased by about two to three times since then. While it is difficult to estimate the amount or percentage of pornographic material relative to other material online, some estimate twelve percent of all websites contain pornography and twenty-five percent of all search engine requests relate to pornography. Neuroscientist Ogi Ogas claims that out of the million most trafficked websites in the world, only 42,337 were sex related, which is only about four percent. He also claims that thirteen percent of web searches are for erotic content.[2]

But how much data is actually being stored and transferred? Due to the content type (pictures and videos, not too much text), porn sites tend to be extremely large. YouPorn, a large pornography site, currently hosts over one hundred terabytes of porn and serves over one hundred million page views per day. That comes out to be 800 gigabytes per second for just one site. Some of the most trafficked porn sites have as many page views per month as Google and Facebook.[3]

While forty-three percent of all Internet users view pornographic material online and seventy-five percent of people "accidentally" have viewed a pornographic site, eighty-one percent of Americans believe federal laws against Internet obscenity should be vigorously enforced.[4] (That's not a typo.) That means a large percentage of the people looking at pornography think it should be illegal.

An activity (usually) filmed behind closed doors, perhaps it is appropriate that this seldom talked about, often hidden industry is a primary tenant in the out of sight world of data centers.

1-Jerry Ropelato, "Internet Pornography Statistics." <http://Internet-filter-review.toptenreviews.com/internet-pornography-statistics.html>

2-Julie Ruvolo, "How Much of the Internet is Actually for Porn?" *Forbes* <http://www.forbes.com/sites/julieruvolo/2011/09/07/how-much-of-the-Internet-is-actually-for-porn/>

3-Andrew Tarantola, "How Much Porn Does the Internet Hold?" *Gizmodo* <http://gizmodo.com/5899327/how-much-porn-does-the-Internet-hold>

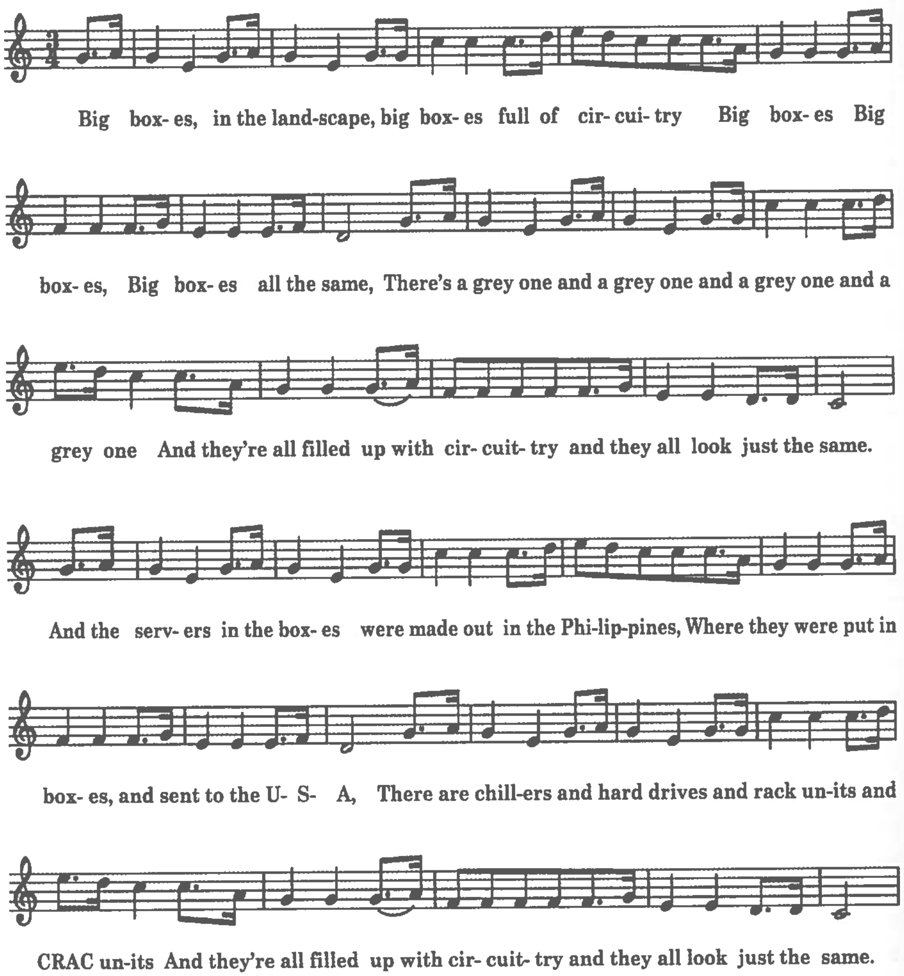
4-"The Internet Porn "Epidemic": By the Numbers." *The Week* <http://theweek.com/article/index/204156/the-Internet-porn-epidemic-by-the-numbers>

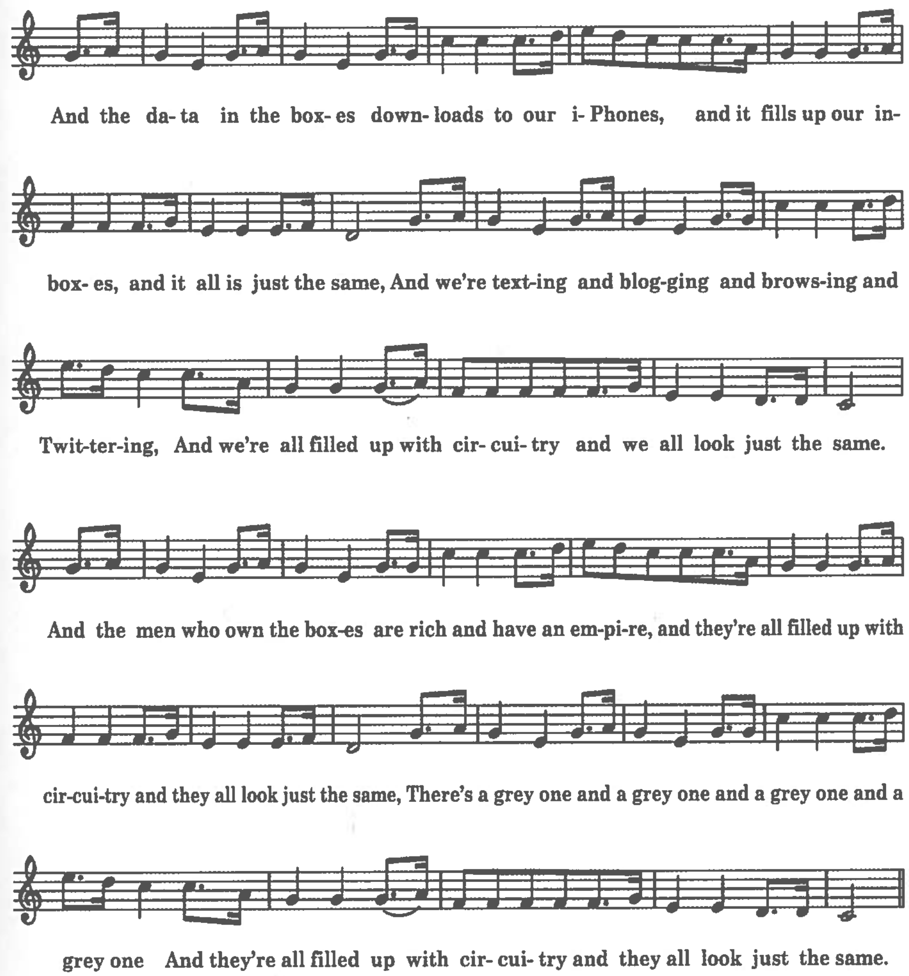
BIG BOXES

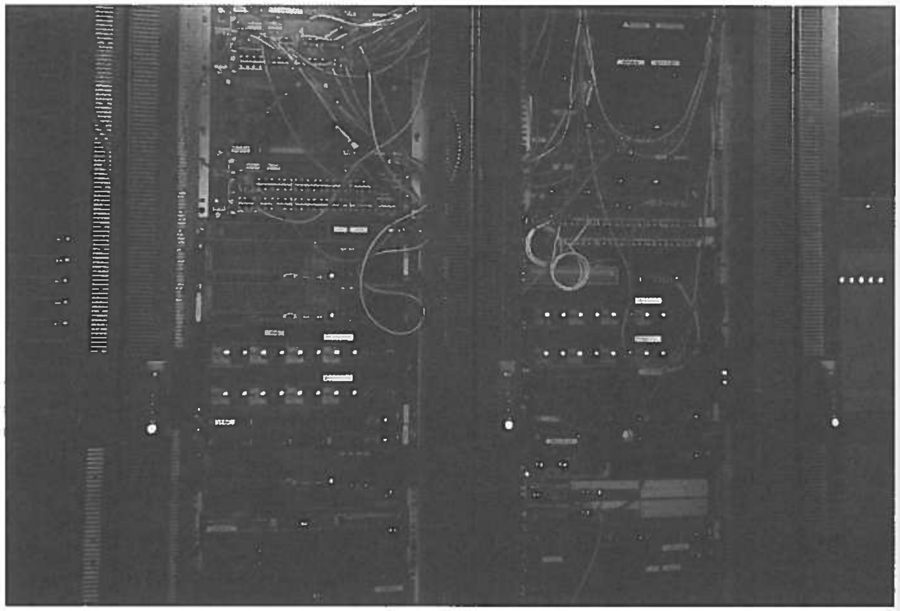
an Ode to the Data Center

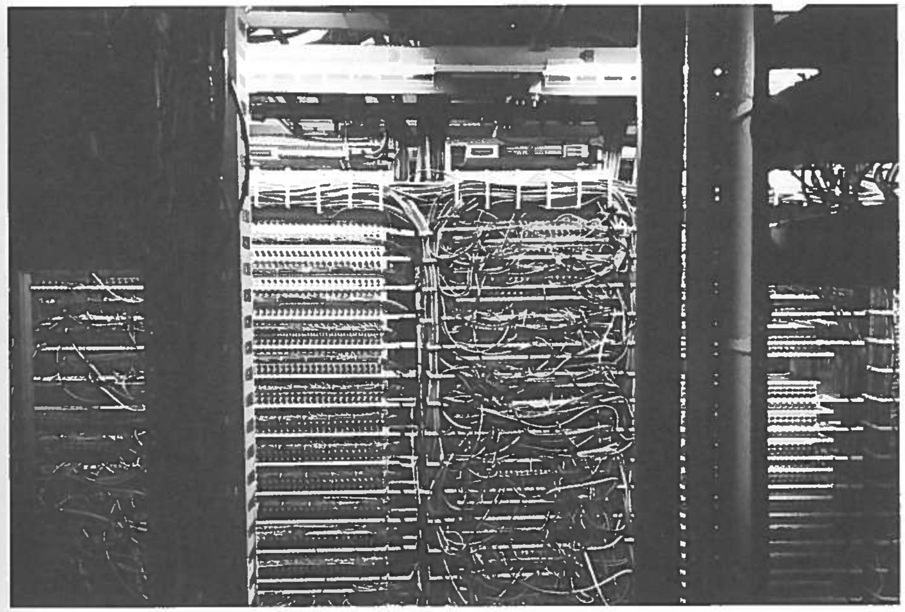
Lyrics by **Ed Ogosta**
(Music adapted from
"Little Boxes" by Malvina Reynolds)

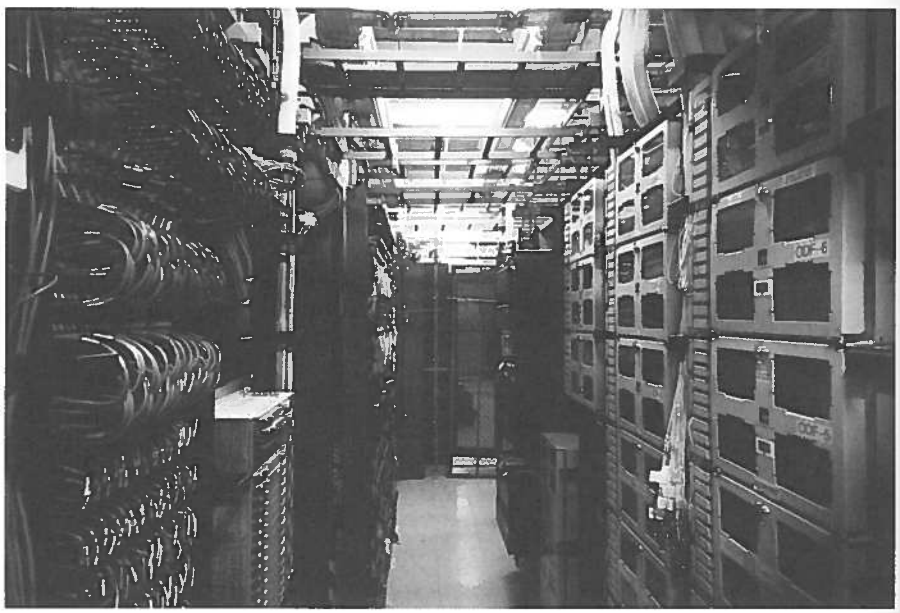
Andante (satirically)

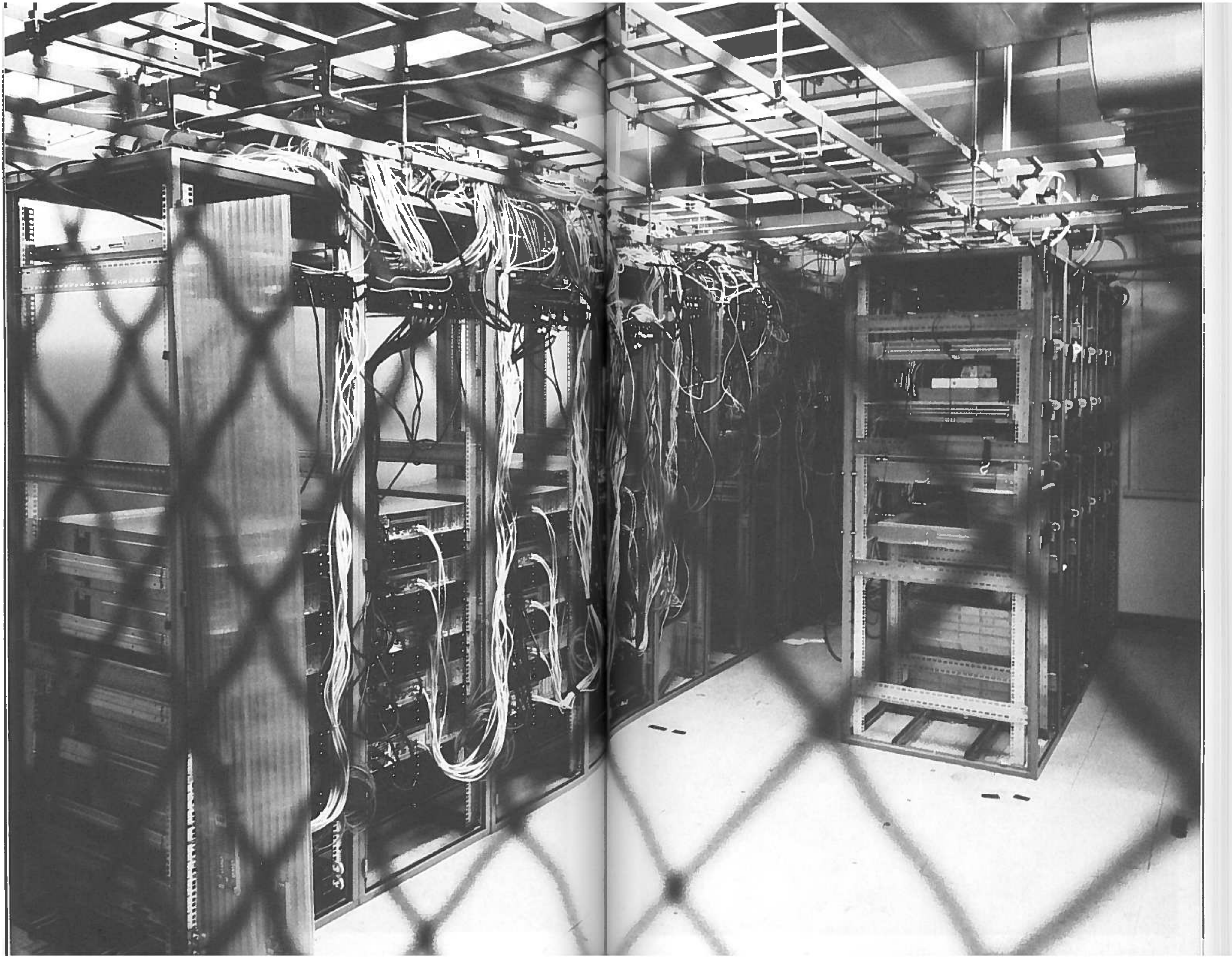


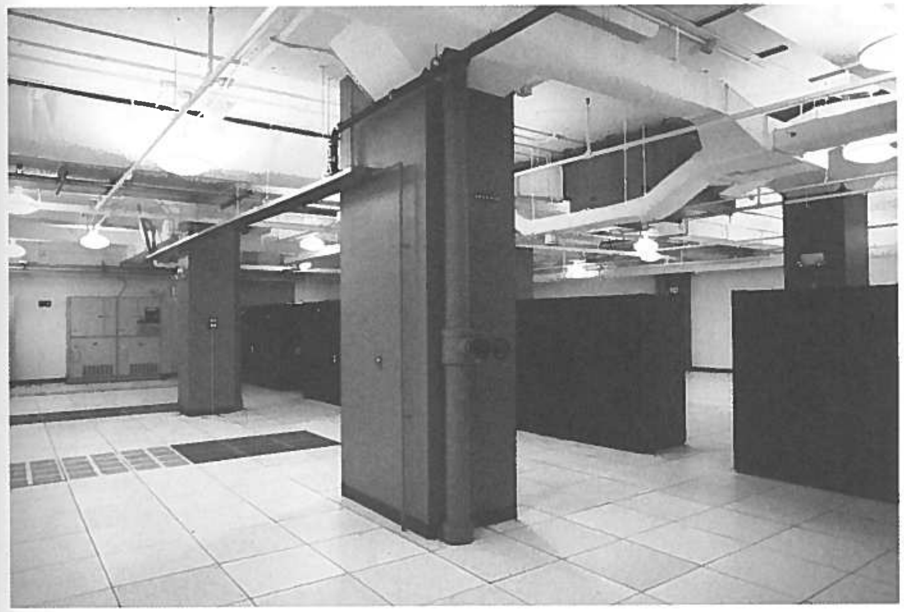


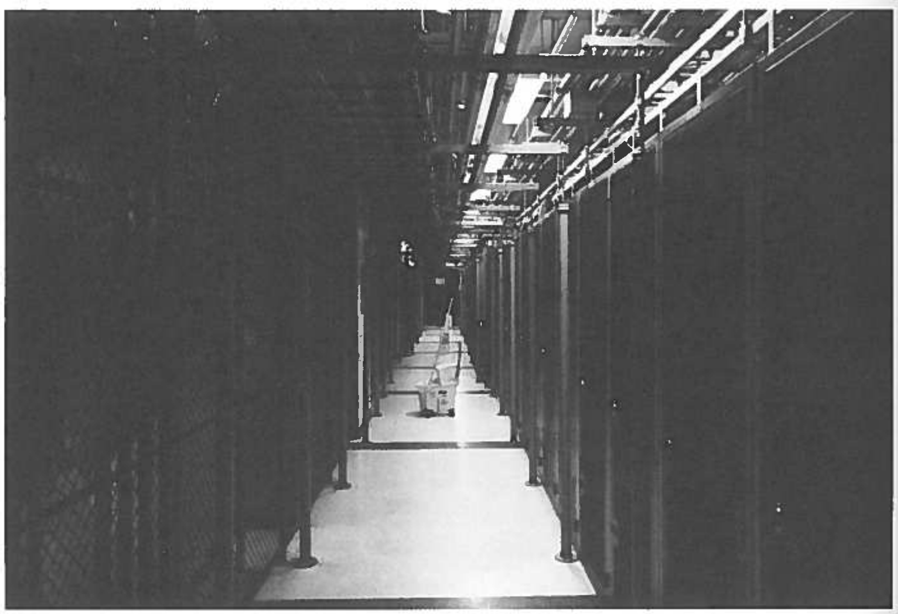


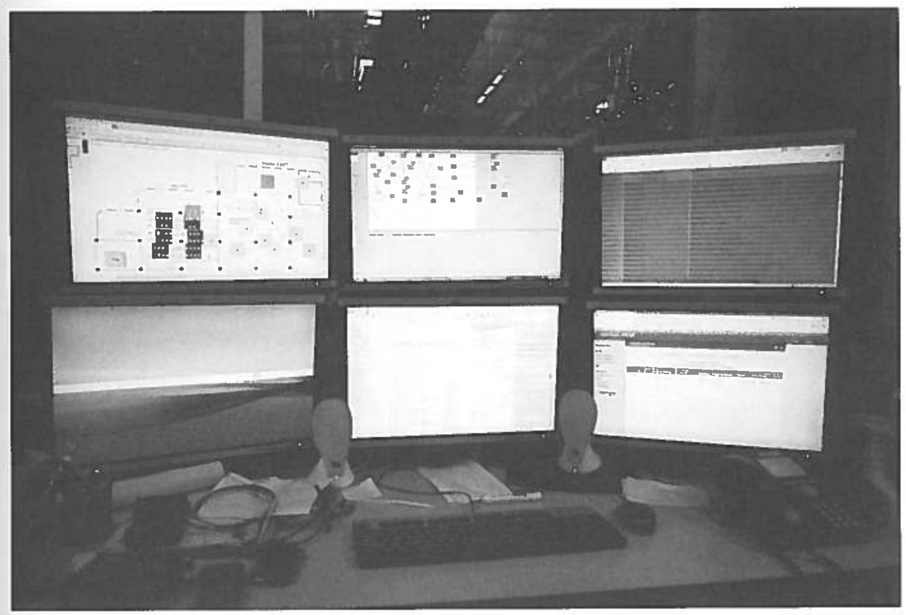


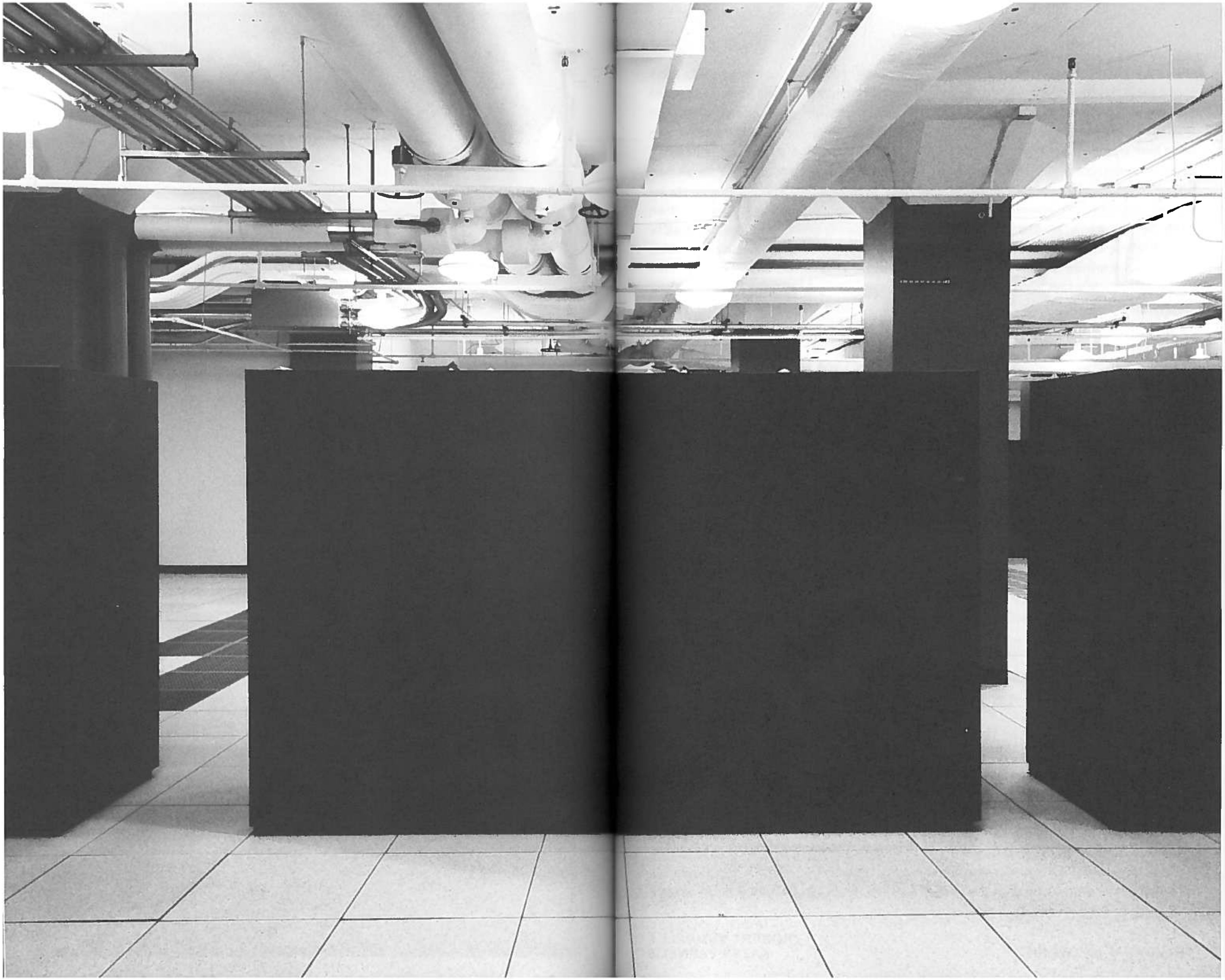

ODF-6
ODF-5



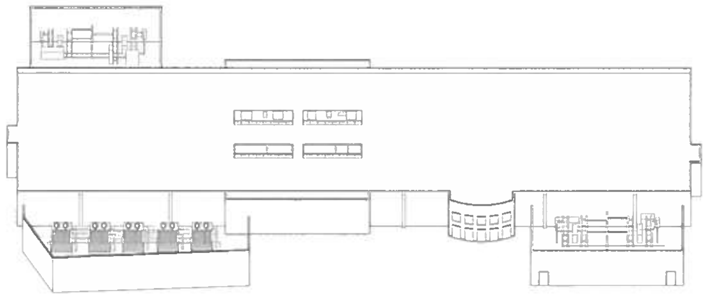




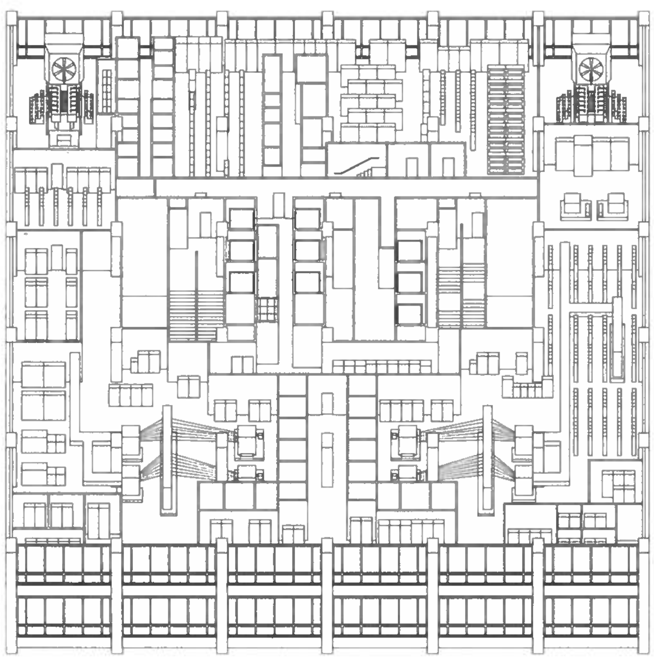




1



2

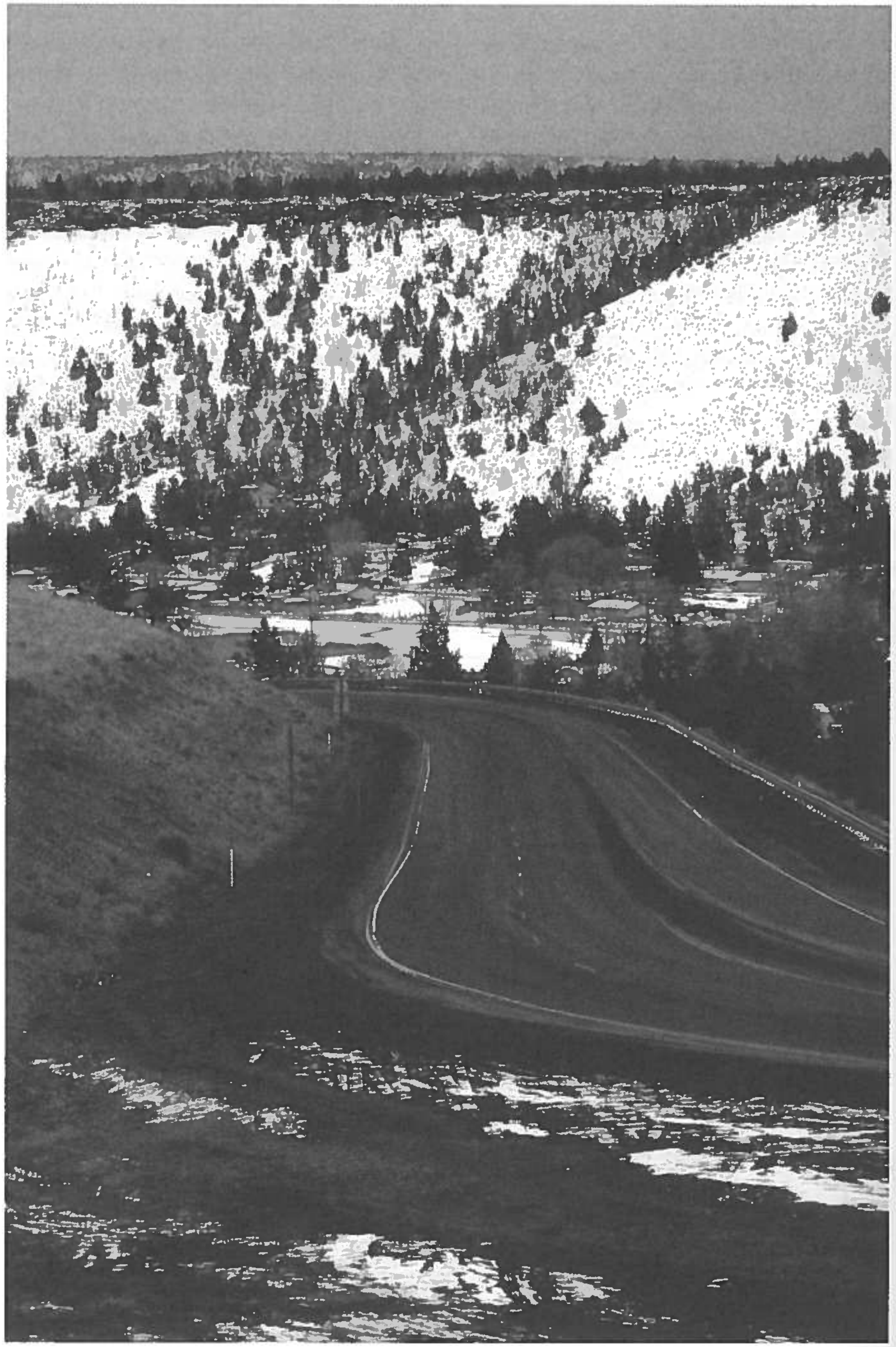


Colocation spaces, in which multiple service providers share a common, highly-connected space, have played a crucial role in the development of the Internet. Throughout the Net's adolescence, telecom hotels like One Wilshire in Los Angeles (2) provided both local and global connections, selling access to nearby links, while simultaneously giving providers the ability to establish direct connections with other providers. As the Internet matured and as mobile network devices became more common, a second type of data center came into prominence, in the form of dedicated rural data storage facilities for companies like Google and Facebook, offering vast memory banks connected to high-speed pipes. Today, however, a new kind of colocation space—ruled by financial concerns—is becoming prominent.

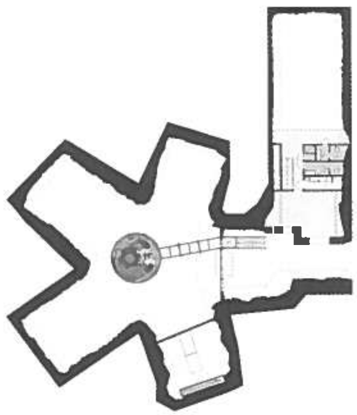
During the last three decades, worldwide investment has grown precipitously. Formerly the reserve of the wealthy, the stock market has been flooded with new liquidity from individuals, investment plans, mutual funds, together with the winnings of the newly rich from developing countries, petrodollars and sovereign wealth funds. With over-investment, profits have fallen and once the dot-com and housing bubbles collapsed, the market was exposed as fundamentally weak. In response, the dominant players in finance abandoned their traditional roles in long-term investment in order to adopt more abstract conditions, extracting returns from fluctuations in the market.

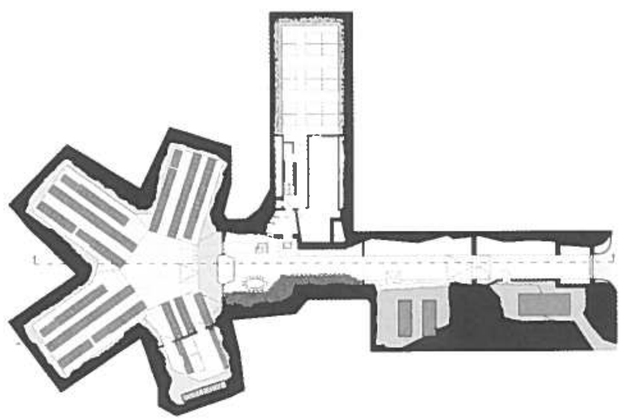
Where One Wilshire was the archetype for the colocation center of the mid-1990s to mid-2000s, NYSE Euronext's Data Center—a 400,000 square foot trading hub in Mahwah, New Jersey (1)—provides today's model for the physical presence of immaterial data. Where One Wilshire served the communications industry, NYSE Euronext services high-speed stock traders, offering colocation space in which clients can make trades at speeds imperceptible to humans. Whereas One Wilshire was a structure in which we all dwelt telematically through our desire to be part of a growing Net, NYSE Euronext services only an elite few and only machines, not individuals. Using algorithms, high-speed traders intervene within the constant flux of the market to profit off of fractions of cents, producing a new alchemy that eschews any need to understand the fundamentals of the financial instruments being traded or even a consistent trend in the market, upward or downward. Instead, it is enough to find ways to exploit the variations at microsecond speeds. With the growth of algorithmic trading, face-to-face trades have been become obsolete and data centers are replacing the familiar trading floors, cementing the idea that human transactions are finally no longer necessary in the newest, more immaterial iteration of the global city.

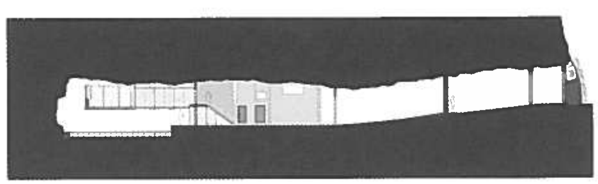
Ninety percent of high-speed trades are cancelled before they are executed, amounting to nothing more than attempts to gauge reactions. The finance industry has thus become a social network—an endless flux of status updates and pokes maintained in real time. These meaningless interactions provide a stimulus to the market. It is this constant fluctuation, not value, that drives the market and with it, society itself. NYSE Euronext becomes the consummate place to check into, even as it stands virtually unoccupied by humans.



Facebook operates its substantial infrastructure primarily from two brand-new, massive data center campuses, in Prineville, Oregon and Forest City, North Carolina. Each is linked to the rest of the world via underground fiber-optic cables, freshly dug to serve the buildings' massive bandwidth needs. In this photo, roadside posts demarcate the cable's path up to the bluff on which the Prineville Data Center sits, overlooking the town of Prineville and Oregon's Ochoco Mountains.







Pionen-White Mountain is a data center built in a former 11,950 square foot Cold War era anti-atomic shelter—an amazing location that is one hundred feet below the granite rocks of the Vita Berg Park, just outside of Stockholm. In 2007-2008, it was converted into a data center by blasting away more than 141,300 cubic feet of solid rock. As the original purpose of the bunker was to be able to withstand a near hit by a hydrogen bomb, this just may be the best protected data center in the world.

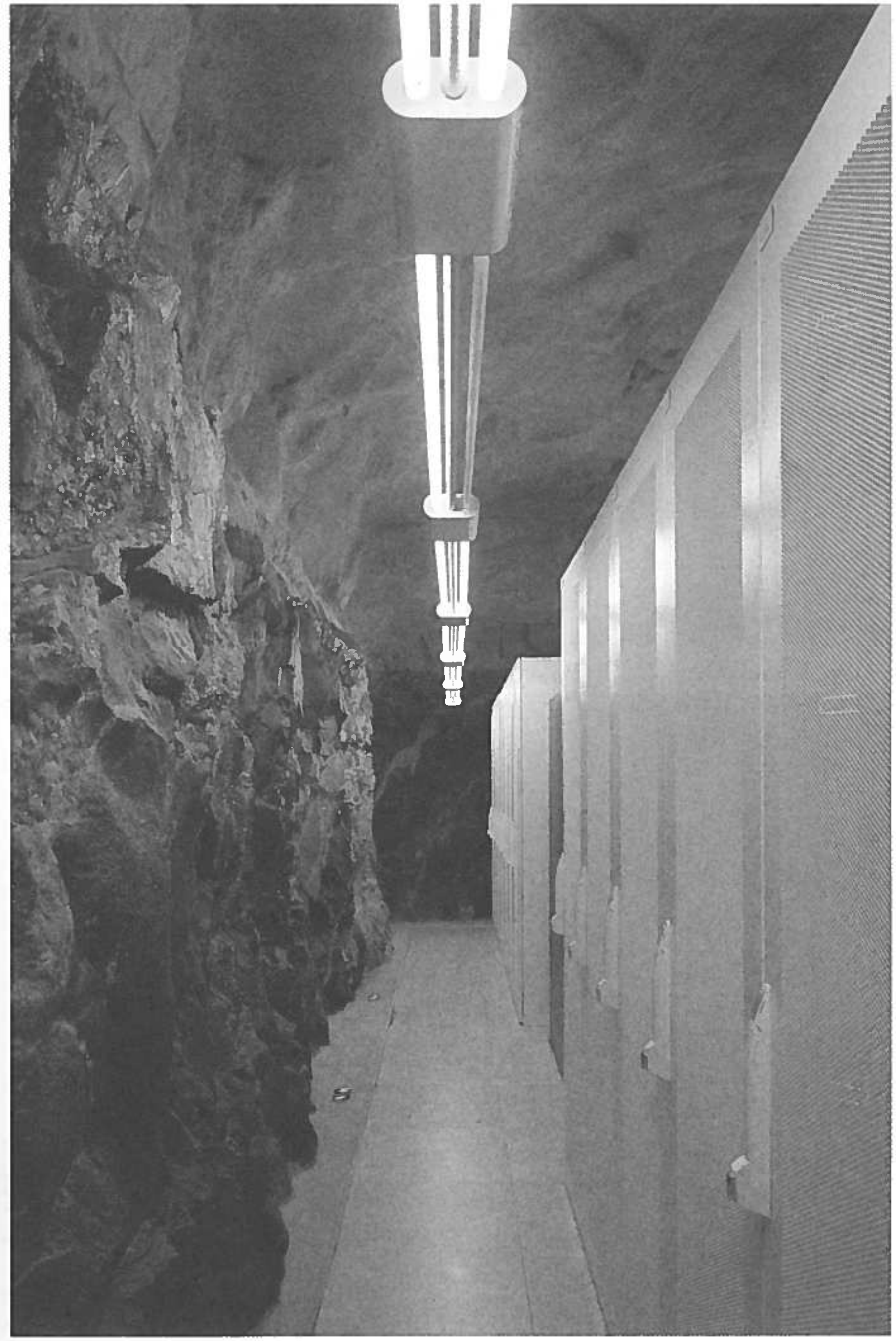
The client is Bahnhof, one of Sweden's largest Internet Service Providers. The company has five data centers throughout Sweden, but Pionen is the largest and it houses the company's network operations center. The facility also acts as a colocation center, renting out server space to other companies.

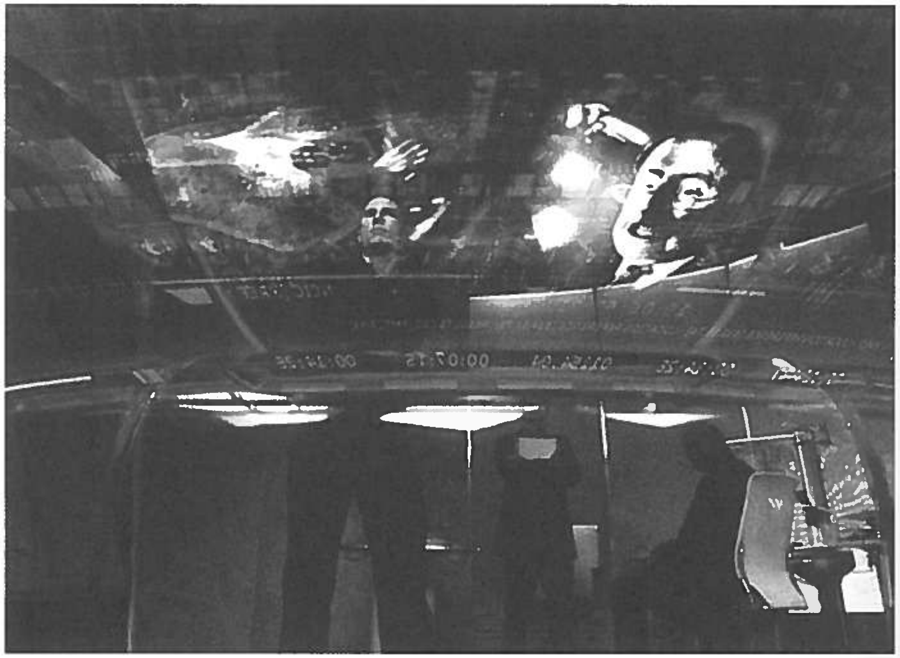
Building a data center one hundred feet below the ground offers several unique design constraints. Backup power is provided by two Maybach MTU diesel engines, which were originally designed for German submarines and produce about 1.5 Megawatts. Cooling is handled by Baltimore Aircoil fans producing enough cooling for several hundred servers. In addition, the network has full redundancy with both fiber optics and extra copper lines with three different physical ways into the mountain.

Unlike most data center designs, the goal was to put humans—not hardware—as the focus. Humans try to acclimate themselves to this foreign world and bring the 'best' elements from earth: light, plants, water and technology. For the employees, Pionen offers simulated daylight, greenhouses, waterfalls, and even a 687-gallon saltwater fish tank. The data center couldn't be a typical hosting center, because of the radical site. Inspiration for the CEO and designer included science fiction movies such as *Logan's Run*, *Silent Running*, *Star Wars*, and James Bond films.

The starting point of the project was to consider the rock as a living organism. There is a strong contrast between rooms where the rock dominates—where human beings are strangers—and rooms where humans have taken over completely. There are strange adjacencies set up by having pristine server racks adjacent to rough granite rock walls and conference rooms hanging within caves. There aren't too many data centers that look like this.

But why are the aesthetics of this data center so important? Again, it goes back to the desire to design for the people working in the space. But in addition, as a colocation facility, there is a need to advertise to clients. Many clients take tours through the space to see the facility. And who wouldn't feel comfortable having their data stored under one hundred feet of granite with James Bond protecting it?





In 2054, according to Steven Spielberg's film *Minority Report*, future crimes will be forecasted and criminals arrested before the act committed. Psychics will gather images and video recordings that document such future events, which, in turn, will be archived and presented to the *PreCrime* Police Department on a transparent, multi-touch, 3D, hologram screen. On this crystal clear surface, past, present and future events will be made transparent and available to the detector's scrutinizing gaze.

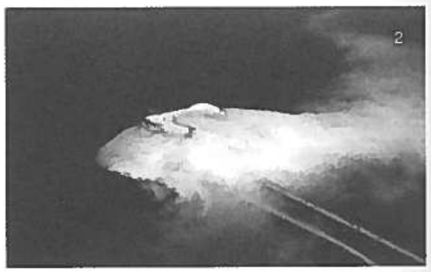
In Spielberg's vision, transparency of information leads to transparent surfaces and, later, architecture. Prompted by such examples, many have argued that developments in telecommunication and electronic technologies are bringing forth a new kind of sleek and light architecture that fits an era characterized by unbounded access to information.

However, Pionen-White Mountain Data Center, the most famous in the world and home to the Wikileaks server, provokes a different kind of architecture—one of enclosure and fortification. Buried one hundred feet underneath Stockholm's White Mountains, and sealed behind a sixteen-inch thick door, its servers occupy the renovated space of Pionen, a former nuclear bunker from the 1970s. Its thick granite walls are covered with vegetation, water fountains, and two ex-submarine engines. With no windows or facades towards the outside world, the architecture of Pionen opposes the logic of transparency associated with the contents it stores.

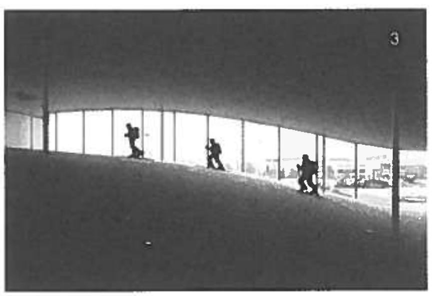
Given a closer look, the architecture of Pionen is more complex. Its lines of fortification, in the context of the twenty-first century, are useless. Its lush green plants seem artificial and its layout—blueprints of which have been exhaustively featured in various design reviews—is anything but a secret. According to Albert France-Lanord, the architect of Pionen-White Mountain Data Center, the references that guided his design came from the 1972 science fiction movie *Silent Running*. Yet, the relation of *Silent Running*'s stage set to a 'real' model is loose. Informed, mostly, by what was believed to be the future of spaceflight in the early 1970s, it vaguely mimics Buckminster Fuller's geodesic domes.

From fake lines of fortification, to *Silent Running*, and then to Buckminster Fuller, the architecture of Pionen is a copy two times removed from a 'real' model. Freed from direct references, it fuses the 'real' and the imagined, concrete and virtual.

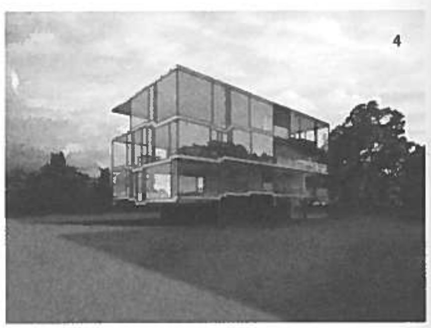
Challenging the bipolar division of transparency and opaqueness, Pionen brings forth a new building typology of data centers—one that embraces the virtual content it stores into its design. Therefore, it differentiates itself from such typological precedents as the cinema and theater hall structures also dedicated to the consumption of virtual material. While the latter ones insist on their 'realness' by means of such elements as the Fourth Wall—that imaginary wall located between the stage and the audience—and maintain conventional seating platforms, the architecture of Pionen oscillates between authentic and fake constructions. In doing so, as a data center, Pionen integrates its peculiar program with form and matter.



2



3



4

DATA CENTERS AND PHYSICAL SPACES: TRACKING THE INVISIBLE

STEFANO CORBO

"In the early '80s we began to live in two kind of cities. One is the city as a material object, which is physically present and supported by physical objects. Beside that, the city as phenomenon is the city that arose with filtered media in societies that had been developed suddenly in '80s. It is the city as information and is a virtual city as event. In this kind of city time and space are not stable."
— Toyo Ito[1]

It's about forty years since the first visions of new cities dominated by intangible connections and data storage centers appeared. In Arata Isozaki's 1972 Computer Aided City, for instance, a huge computation center determined the functioning of various urban activities and everyday life.

Such proposals—the result of technological optimism and youthful radicalism—have proved prophetic. After four decades, networks and hubs characterize the spaces we live in and determine our social, psychological and perceptive behaviors. In order to survive the progressive dematerialization of processes and relationships, architecture can only re-invent itself as the infrastructure of invisible data: not in the *machine à habiter* Modernist way, or as an abstract support for human activities, but as a dispositive that is able to detect, capture, and re-interpret the invisible data that characterizes contemporary society.

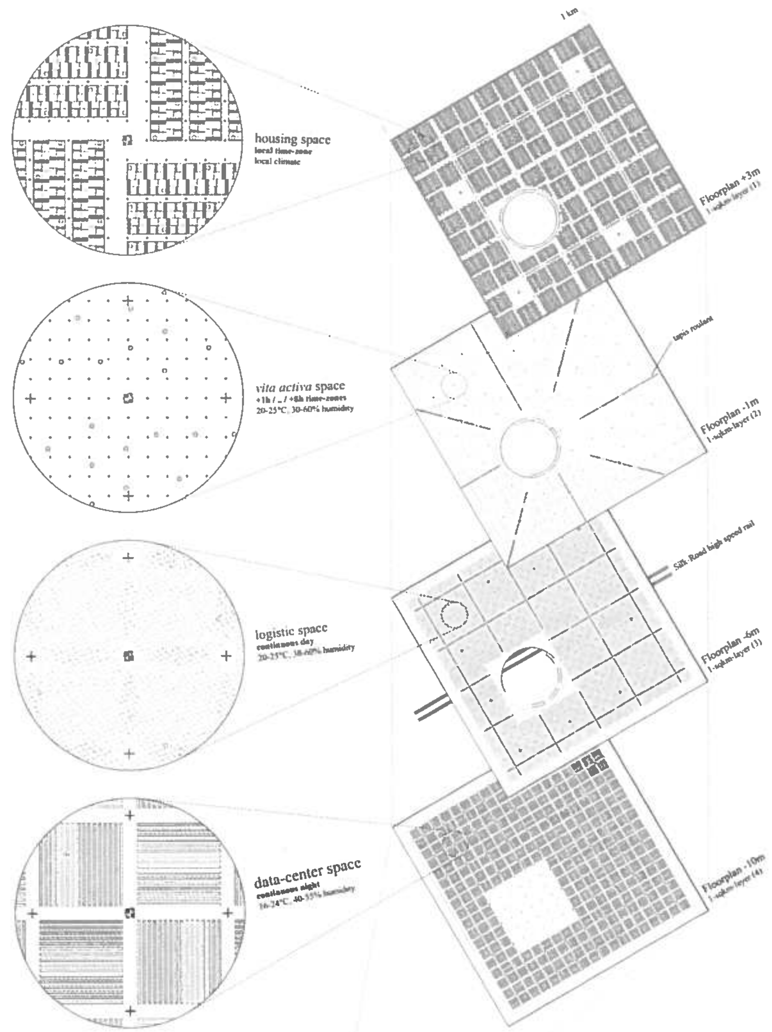
Architecture as infrastructure would therefore serve as the diagrammatic representation of these processes, canalizing them into the design of real spaces. Through the use of these "weak" infrastructures, architecture's task is to track the invisible and uncover the network of elements, points and liquid relationships that characterize our society (1). Vision Machines, Knowledge Apparatus, Convective Architectures are just some examples of how the discipline of architecture has tried to address this need to integrate the intangible conditions of the contemporary city with the concrete design of physical spaces.

In Diller Scofidio's vision infrastructures (2), the eye is the primary shaper of space and psychic organizations. Every room or building is revealed as a machine that distributes visibilities and invisibilities, trapping the spectator in a state of total self-perception. Toyo Ito or SANAA's Knowledge Apparatus (3) intends to order the intellectual environment by introducing information technologies and innovative forms of data control. These open and fluid structures try to perform services that are not limited to the museum, but will continue to evolve and share the network. In Philippe Rahm's thermal landscapes (4), the user wanders around as if in a natural landscape, looking for specific thermal qualities and different temperatures related to the moment of the day or the season.

These approaches represent different attitudes and distinct methods, but share a common starting point: architecture can no longer deal with those elements of our existence that are strictly material and physically visible, but through its design, must let the hidden and invisible forces of our everyday relationships emerge.

1-Toyo Ito, Escritos, Colección de Arquitectura, n.41, Cajamurcia, Murcia, 2000.

1
2
3
4



The 2010 World Expo in Shanghai prompted a geopolitical investigation into the possibility of developing a high-speed rail line following the ancient Silk Road. Such titanic infrastructure would potentially challenge all other global logistical routes. To function, the proposed 8,000 kilometer high-speed rail line would require urban centers along its length to stabilize its use, as well as a parallel dedicated data network to sustain the complexity of its real-time management and to coordinate the myriad energy plants necessary to power the system.

We therefore designed a one square kilometer footprint settlement that could act as a discrete unit of development along the Silk Road, capable of disentangling the complexity of this intercontinental infrastructure. Considering the multi-scalar relations at stake, the settlement has been organized into four superimposed one square kilometer layers: (1) housing with a local circadian rhythm of light and temperature; (2) vita-activa zones for work and commercial activity tuned with the circadian rhythm of each time zone crossed; (3) a twenty-four hour logistical operation zone with continuous day conditions and an environment designed for human comfort; and (4) a twenty-four hour data center with continuous night conditions optimized for IT operations.

What are the capacities of such a data center? Considering that one server rack today can reach 12.3 TeraFLOPS and requires almost 4.5 square meter to be managed, the settlement's computational power would be 2,678,000 TeraFLOPS, corresponding to 133 supercomputers or 535,600 high-end PC's. In theoretical terms, this is 1/400 of the computing power required to perform full weather modeling. With 400 settlements (one settlement every twenty kilometers of the Silk-Road) a computational infrastructure powerful enough to permit extremely long-term, precise forecasting of global weather patterns might actually become possible. This capability could, in turn, potentially allow humans to model the impact of our actions on the global climate, and to modify these actions accordingly.

The one square kilometer data center would require 1,785 MegaWatt to operate, since each MegaWatt today permits 1,500 TeraFLOPS (hardware cooling included). This could only be consistently supplied by a large-scale atomic plant, which today can produce 2,000 MegaWatts. Conversely, the 40,000 inhabitants themselves require only 1/70 of the energy demanded by the data center. To support the inhabitants' energy usage, photovoltaic panels on the roofs of the housing layer could provide twice the energy necessary for domestic activities. The integration of photovoltaic and atomic energy production would essentially satisfy all the settlement's energy requirements, allowing its residents to focus solely on managing the complexity of the new Silk Road.

"We are in the age of the simultaneous, of juxtaposition, the near and far, the side by side and the scattered."
—Michel Foucault,
Of Other Spaces:
Utopias and Heterotopias,
1967

We are living on and around the boundary between physical and virtual data space. Today, architecture is not only limited to physical/tectonic space, but it also points to virtual data space. While our spatial boundary between physical and virtual space is being overcome, for example, through the integration of virtual shopping spaces with other types of physical spaces, our understanding and perception of space is also being changed with simultaneous and parallel experiences, where physical and virtual spaces co-exist.

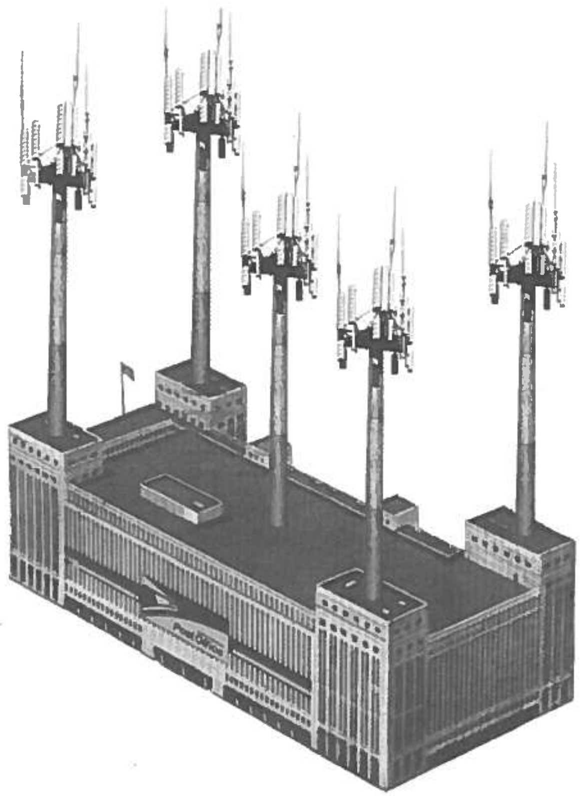
Large corporations are assembling enormous buildings to process the world's digital information, paradoxically triggering the consumption of relatively cheap land and wasting energy resources. New methods of synthesizing data centers with their dense suburban centers allow the reuse of otherwise wasted energy while calling awareness to the Internet's monumental physical footprint and output of waste heat.

Two projects that best exemplify this are the Academica data center in Helsinki, Finland, and the Telehouse West London Docklands data center in London, England. Both of these projects use innovative solutions to reduce the amount of energy they consume. Initially, they both extract the heat produced by the servers to power the surrounding neighborhoods.

Taking cues from these strategies, an aggregated suburban landscape in the city could capitalize on the trend of enormous data center expansion and simultaneously subvert the trend of suburban sprawl.

Ranking first in the United States for data center placement, Iowa faces growing suburbanization at the sacrifice of much of its most valuable virgin agricultural land. In the midst of this ubiquitous expansion of suburban sprawl, we arrive at the critical moment to end this recklessness by considering the role of emerging data centers.

A solution could be a centralized public platform with mixed-use developments as illustrated in the birds-eye image to the left. The new mile-long data center could be used as a "microclimate platform" for cultural activities and a contemporary public identity expressive of a new age of industrialization. For example, capturing waste heat provides the opportunity for a heated sidewalk in the winter, promoting the further connection of adjacent neighborhoods. With the savings the city may anticipate from infrastructural growth and maintenance, the advent and financing of public space in the suburbs will be possible. Such a solution simultaneously promotes the expeditious construction of data centers as well as the re-centralization of a public platform for a polycentric city.



NICHOLAS MCDERMOTT,
NATHAN RICH &
BENJAMIN SMOOT

SYSTEM UPGRADE

While data centers might represent the potential for a new building type, as a societal tool they represent an evolution, an upgrade. The technique is novel, but the service, long distance data and voice communication, is not. Data had been moving around by physical transportation of the written word, telegraph, telephone, and wireless signal for centuries before computer networks started performing a similar task.

It is the case, however, that the Internet has become so efficient, and has created so many new types of communication, that earlier methods are now losing ground. No contemporary method of communication has suffered more in the face of the new competition than the handwritten personal letter, fewer of which are being mailed every year.

As goes the letter, so goes the postal service. The decline of the former has led to serious talk in the United States about the possibility of a radically reconfigured mail system. For one thing, the postal service is thought to have too much physical presence, in the form of employees, vehicles, and significantly real estate. Given the trend and current government forecasts, consolidation and selective elimination of these physical resources seems inevitable. But the hard times at the United States Postal Service (USPS) have been boom times for new networks, and data infrastructure is expected to grow significantly in coming years.

So here's the situation: there are two competitive services, one in need of serious dematerialization and the other with increasing demands for more physical space. It's an opportunity for a total system upgrade, a nationwide one-to-one adaptive reuse, and the benefits could be significant.

For example, while market forces currently drive decisions about where broadband access is located, the USPS has traditionally operated a minimum of one post office in each zip code. This means that there is an even distribution of publicly owned postal space that, if maintained, could be converted to a new type of data space. The death of one public typology—the community-centered post office—could foster the growth of a new one—the close proximity high speed access point. In the process, a feature which is now a financial handicap for the USPS—physical ubiquity—could become an asset; an adaptive reuse campaign focused on the postal system would have the scale and transformative potential of successful past infrastructure projects like the interstate system, while creating new social possibilities for an evolving typology.

By integrating access points and communication tools within reinvigorated public institutions, community connectedness would increase at a time when it is being threatened by the liquidation of a traditional public space. It would create a new physical and democratic interface with network and data-based resources. Data space would become less proprietary, more open-source, and the creative potential of networks, now expressed most powerfully within the virtual confines of social websites and financial markets, might find its way into the physical world.