Century-Scale Smart Infrastructure

Dhananjay Jagtap
dtjagtap@eng.ucsd.edu
University of California, San Diego

Nishant Bhaskar
nibhaska@eng.ucsd.edu
University of California, San Diego

Pat Pannuto
ppannuto@ucsd.edu
University of California, San Diego

ABSTRACT

On average, wireless electronics devices are replaced every 50 months. On average, a bridge is replaced every 50 years. As we begin to imagine integrating electronics and intelligence into the built environment, we need to begin to think about electronic devices and systems on infrastructure timelines. This is not to say that every individual electronic device can, will, or should last for decades, but much like the ship of Theseus, the system that defines emerging Smart Cities will have a lifetime reaching into the century-scale. In this paper, we contemplate what the devices, gateways, network architectures, and their management might look like for a system designed to operate for decades. The result is a mixture of actionable insights for today and research questions for tomorrow, which culminates in the commencement of a 50-year experiment designed to see how long energy-harvesting sensors, without the implicit lifetime of batteries, can remain viable without human attention or intervention.

CCS CONCEPTS

• Computer systems organization; • Networks;

ACM Reference Format:

1 INTRODUCTION

In the United States, a wall socket today provides the same 110–120 V, 60 Hz signal\(^1\) as it has since the establishment of a national grid system. While much of the original telephone network has evolved, “last chain”\(^2\) links in homes are still copper, and modern VoIP modems still include circuitry to support pulse dialing for rotary phones\(^7\). Wired power and communication infrastructure have supported billions of devices for nearly a century.

While there likely are some still-operational devices approaching 100 years of service, the power of the infrastructure promise is not in supporting the long tail of long-lived devices but in the promise that any device could be a long-lived device. It invites investment for functional obsolescence, “if it ain’t broke, don’t fix it,” which maximizes device utility and return on investment over time.

In contrast, wireless electronic devices are more often subject to technical obsolescence, when a newer, better, or more capable device supplants existing technology.\(^3\) Sometimes this is driven by capabilities of the device itself, e.g., a more powerful processor, but sometimes it is driven by external changes in technology, such as the upgrade of a wireless scale that supports only 802.11b to improve the performance of a home network\(^11\). When a device relies on infrastructure, the device’s lifetime is the shorter of the lifetime of the device itself and the lifetime of the infrastructure it relies on.

Replacing a modest number of otherwise functional personal devices due to technical obsolescence is annoying but manageable. Replacing a city’s worth of devices is intractable. Consider the scale of Los Angeles, where there are over 320,000 utility poles in service\(^29\), 61,315 intersections\(^8\), and 210,000 streetlights\(^9\)—three common targets for monitoring sensors. If critical communication infrastructure disappeared, at a very generous 20 minute total replacement (including travel) time per device, recovering the deployment would require nearly 200,000 person-hours of labor alone.

This large volume of infrastructure does see maintenance, repair, and upgrades, which includes the opportunity for economical sensor replacement. However, it is important to recall that Los Angeles was not built in a day. Instead of replacing or upgrading one sensor type en masse, infrastructure projects operate in geographical batches to keep costs down—one project repaves a block, installs its traffic sensors, and replaces its streetlights.

\(^1\)Grossly simplifying, the goal was always 120 V, but early end devices saw voltages closer to 110 V for various technical reasons. Distribution has improved such that today most endpoints realize 120 V\(\pm 5\)%.

\(^2\)chain; noun: A surveyor’s unit equal to roughly 20 m.

\(^3\)They may also be subject to (particularly in consumer electronics) style obsolescence, where otherwise functional devices are replaced for reasons of personal taste, and (sometimes controversially) planned obsolescence, where a device’s lifetime is limited by the manufacturer, either via components designed to fail or explicit software lockouts\(^23\).
Like one view of the Ship of Thesues, the lifetime of a sensing system then is the aggregate lifetime of all of its devices across all their deployments. Constituent device lifetimes are pipelined, where some 15-year sensors are 10 years into their service life while others are being freshly deployed. In any one build-out site, some deployments are replacing their sensors with state-of-the-art technologies, while others are deploying legacy devices to keep costs down or lessen operational heterogeneity. The takeaway is that even if it is unlikely for any one device to last multiple decades, it is both reasonable and likely for municipal-scale systems to last for decades.

With the more immediately-likely path to decades-long systems laid out, it is worth also spending some time considering the possibility of widespread, decades-long, individual devices. There are already "Low Volume Complex Electronic Systems," such as military technology and satellite hardware, whose operational lifetime exceeds 50 years [30]. Building on the failures and maintenance lessons from these systems, might we be able to realize general-purpose devices with timelines akin to the physical systems they are monitoring?

Recent advances in the energy harvesting and batteryless sensing community have made significant headway to addressing this challenge. So-called "Ambient Batteries" find stable, battery-like energy sources that could (theoretically) power deployed systems for a decades or more [20, 21]. What wireless technology should be employed by a sensor physically embedded in the concrete matrix of a road (median service life of 25 years [37]) or a bridge (median service life of 50 years [31]) that reports on the actual concrete health [34] and powers itself—for literally as long as the structure lasts—off of the corrosion of the embedded rebar?

While there may remain some hurdles to the realization of such long-lived hardware today, if one accepts the premise that there is value in embedding sensing in infrastructure, then either the lifetime of embedded sensors must rise to match the lifetime of infrastructure or the infrastructure lifetime will fall to match that of the sensors. And sensors are only useful if they are able to communicate their data.

In sum, this paper looks at the challenges of pushing Internet of Things systems, particularly edge sensors, from a paradigm of technical obsolescence to functional obsolescence. To that end, we close the paper with the commencement of an experiment that challenges assumptions of the real-world lifetimes of today’s devices. Conventional wisdom holds that components such as batteries, electrolytic capacitors, or even PCB substrates will hold the mean lifetime of a device to around 10–15 years [19, 22]. Energy-harvesting devices require no batteries, however, and the same manufacturing processes and circuit design points that make systems low-power also make them more robust to long-term failures. We commence a 50-year experiment, one that suggests that if we put the effort into maintaining the supporting infrastructure, we may already have sensors beginning their century-scale operational lifetime today.

2 WHAT IS SMART INFRASTRUCTURE, WHO WANTS IT, AND WHY?

Smart infrastructure can be any physical infrastructure that responds intelligently to environmental changes or user demands. This involves deploying individual sensors nodes, actuators, and networking support within the built environment. One of the most widespread examples today is advanced metering infrastructure, which enables two-way communication between utilities and customers.

When, and for whom, is it worth it to pay the costs to increase the reach of the digital infrastructure into the physical world, however? For smart cities, deployment is intended to improve the services that the city provides to its citizens. One city that invested is San Diego, which installed 8,000 smart LEDs with 3,300 sensors on top of streetlights [35]. Each node consists of video, acoustic, vibration, magnetic, and environmental sensors. Deployed applications that process and use this data track parking availability, support law enforcement, and capture longitudinal environmental trends.

Smart cities deployments can count many successes today. They provided disaster recovery and management support in Chattanooga, TN [2], flooding prevention in Calgary, Canada [33], and reduced overflow of trash bins in Seoul by 66% and cost of waste collection by 83% [32] Some exhibit more mixed success. The San Diego deployment is changing backhaul networks [4] and is embroiled in legal disputes over the ownership and recording of video data [10, 28].

From the existing build-outs, we can learn a few key trends. First, there can be monetary benefits from smart infrastructure. Second, the cost for deployment for even for a few thousand sensors can range into millions of dollars. Right now, in cities where such networks are implemented, the numbers of nodes usually range from 500-5000. For these modest numbers of devices, operators predict lifetimes of 2-7 years until the system is upgraded [4, 6]. This limits systems to settings that are comparatively easy to access and replace.

If we want to push to deployments of ten thousand, ten million, or even billions of devices, then the time and cost effort will scale up rapidly as well. Yet there is motivation to do so as the success of an IoT application is tied to the scale of the network. Instrumenting one intersection will not give city planners an accurate picture of the overall city traffic. Air pollution is highly localized, and requires measurement at city-block granularity [27]. To realize smart infrastructure, we need to make the deployment of sensing devices as easy and reliable as plugging in a telephone or appliance today—only now we must do so without physical wires and plugs.
3 COMPONENTS OF SMART INFRASTRUCTURE

What should a utility (if the utility model even fits) for the IoT look like? How it may differ from existing (non-utility) wireless communication infrastructure? The goal of such infrastructure is to allow private, general public, and municipal entities to realize more economical and reliable deployments. Once, New York City was a grid of grids, with multiple competing electric delivery systems all providing the same fundamental service. Communication technology is more nuanced than power, one size likely will not fit all, but it is distinctly possibly that a few sizes can fit most. History shows that the right, standardized infrastructure will ease the effort and cost of deployment and speed up the scale of connected devices. To understand what such infrastructure for Smart Infrastructure may look like, we break the system down to broad hierarchical levels, shown in Figure 1.

3.1 Edge Devices

In smart infrastructure, devices are the most numerous and least accessible. They are deployed in the real world, are physically embedded in sense points of interest, and are geographically diverse. As the quantity and diversity of devices grow, their accessibility and maintainability falls. Ultimately, there are a finite number of person-hours available for the maintenance and upkeep of sensing systems; as the number of devices grows, the available hours per device falls. The more self-sufficient and deployable devices become, the more we will be able to scale up their deployment and the reach of smart infrastructure.

The capabilities of a device will be constrained by its deployment context and application requirements. As a consequence, devices will show a great degree of heterogeneity. One common thread, however, is that devices will need some means—most often wireless—of networking. The varying resource constraints and application requirements will result in devices with a diverse, but countable, variety of wireless technologies, each of which will require supporting gateways to connect to the network at large.

Once a device relies on a gateway for communication, however, the lifetime of the device is limited by the lifetime and availability of its gateway. One key lesson from early sensor networks research is that even severely resource-constrained systems can use standards compliant (i.e. IP-based) networking [17]. Devices can de-risk their deployment then by relying on the presence of some gateway capable of translating their wireless signal, but eschewing protocols that require authentication to a specific gateway.

Takeaway: To enable scale, individual devices should expect no human attention during their operational lifetime. Takeaway: Devices should rely on properties of infrastructure, but not specific instances of infrastructure.

---

Figure 1: Deployment Hierarchy. Smart devices rely on one or two gateways, while gateways may support thousands of devices. Similarly, individual gateways rely on one or two backhaul technologies, which backhaul infrastructure may support thousands of gateways. The further up the hierarchy one travels, the more devices there are that are reliant on the stability and reliability of the provided interface. Complementing this, when moving down the hierarchy, stable standard interfaces provided by layers of infrastructure enable heterogeneous, unplanned deployment of a wide variety of devices.
3.2 Gateways

Gateways provide connectivity for devices. An important property of the gateway layer then is to maximize geographical coverage. Unfortunately commercial interests of device manufacturers are often at odds with achieving practical coverage. A major deterrent is the lack of interoperability among commercial wireless sensor products. Manufacturers often lock down their software ecosystem, so that their sensors can only work with their specific gateways [3, 38]. Consequently, today’s cities end up containing several ad-hoc wireless systems that are redundant (e.g. co-located 802.15.4 gateways that serve devices from different manufacturers).

Fundamentally, a gateway must route packets between devices and backends. Beyond this, gateways may also need to support different communication channels depending on the management or security needs of the different devices and applications. A gateway collecting connectionless data from transmit-only monitoring sensors may only need to forward data (possibly while minding a blocklist of known-bad devices). In a more advanced case, a gateway unit inside a traffic light controller would also need to maintain connection keys for establishing secure channels with all the sensor and actuator units if local, closed-loop control were desired.

Finally, like all components in the sensor network, the gateway layer must allow for upgradability. This can be achieved through a well-defined commissioning process. The process should allow newer gateways to establish links with the backhaul using secure mechanisms similar to those used for home router commissioning. Additionally, when replacing existing gateway units, we can have a process in place to utilize the outgoing gateway as a trusted third party [36] for easy migration of existing connected devices.

Takeaway: Gateways should primarily act only as routers, and defer decision-making to other system components. Takeaway: Connectivity from gateway deployment can be increased, if gateways provide coverage to all devices regardless of the manufacturer and with the existence of a reliable backhaul infrastructure.

3.3 Network & Backhaul

The backhaul network connects gateways to the cloud. For this discussion, we are primarily concerned with the physical link between gateways deployed broadly in specific physical locations throughout the world and the internet at large. The considerations of capacity and cost play a decisive factor in deciding whether the communication should be wired or wireless. This network needs to guarantee that it will be able to sustain copious amounts of data with sufficient reliability.

3.3.1 Fiber (Wired). Fiber optics support large bandwidth. Furthermore, fiber optic cable capacity depends more on the end transceiver equipment than the actual fiber itself. With improving multiplexing technologies, the capacity of already laid fiber will only go on increasing over time.

Fiber serves as the backbone for real-world smart city applications today. One example is San Leandro, CA where the backhaul communication from all the gateways is based on a fiber optic network [26]. Barcelona, which is one of the most digitally integrated cities, uses an extensive 500 km fiber optic cable network for communication. Remarkably, most of this urban fiber network was more than 30 years old by the time Barcelona started implementing its IoT project [1].

The major chunk of the capital cost of laying optical fiber is digging the physical trenches. For municipalities, this cost can be easily amortized by coordinating the deployment of cables with other maintenance works relating to roads, power, or other public works—indeed, traditional commercial service providers already do this for their fiber deployments. The large bandwidth of fiber permits use beyond supporting deployments, and thus deployment costs can by further distributed in support of alternative services such as community WiFi hot-spots. In San Leandro and Barcelona, they use their networks to provide internet to businesses and homes which has provided another platform for a possible revenue model.

3.3.2 Cellular (Wireless). Using extant cellular technologies like 2G/3G/4G and now maybe 5G for the backhaul communication is easier to implement in most cases as there no infrastructure that needs to be laid. Newer gateways can be added using over the air activation easily. In the long term, however, the operational costs of subscription from service providers becomes expensive. One large cellular-backed smart city deployment, the city of San Diego, is planning a transition to lower cost wired options such as fiber or Ethernet to replace their current 3G/4G communication [4].

Beyond cost, a major risk with subscribing to a cellular backhaul is that no operator guarantees service periods as long as 50 years. Neither, it should be noted, does a wired provider, but wires do provide the assurance that they generally will not go anywhere. Even if the other end of a wire were to go dark, the bulk of the infrastructure cost has been paid, and replacement service could be made available at comparatively manageable cost. In contrast, because spectrum is a very limited and very expensive resource, municipalities cannot easily deploy their own replacement cellular backhaul in the future.

3.3.3 Ownership Models. The backhaul is what enables deployment of future smart city networks. Therefore, the question of who should deploy, own, and manage this piece of infrastructure becomes very important. Detractors of municipal-owned infrastructure often tout the lack of technical knowledge of cities and high sunk costs as possible reasons why backhaul fiber optic/wireless networks should
be handled only by private companies. Empirical evidence from deployments and deployment reports in cities across the US refute this, however (Santa Monica, CA [24], Chattanooga, TN [12], Martin County, FL [15], Mount Vernon, WA [13], and Chanute, KS [14] to name a few). Cities like Santa Monica and Mount Vernon run a complete fiber deployment without even a municipal electric utility. The city of Chanute, KS has been running their own 25 Gbps WiMAX network in addition to fiber, for over 10 years now. And as evidence that small cities should not shy away, Chanute is a city of only 9,000 residents that employs just 2 staff to run their fiber network, which is currently profitable [14].

In addition, city ownership of backbone networks allows local control of end applications and the lifetime of their support. Middling service for institutional networks (i.e. those that connect schools, municipal buildings, libraries) are a classic example of why this is important. Cities like Santa Monica and Martin County originally relied on cable/telecom companies for managing these networks. Private companies provide discounted or free service to municipal needs in exchange for access to public right-of-ways for their infrastructure. Unfortunately, once infrastructure is deployed, municipalities have no direct control over network operation or performance, and their priorities and needs are often underserved, particularly once compared to the service provided by later-deployed municipally owned networks [15, 24].

Takeaway: Backhauls must provide reliability and service guarantees that last or exceed the time that would be required for users to replace them.

Takeaway: Backhaul maintenance is not actually that hard, and even small-scale municipalities should not shy away from deploying their own infrastructure.

### 3.4 What makes Century-Scale?

While some of the principles above may seem straightforward, long lived examples of digital infrastructure are rare today. We re-emphasize that it is the aggregate of end applications which are century-scale. Any one application can fail because devices go offline, the gateways they rely on cease operation, or the backhaul that the gateways rely on cuts service. When devices fail, the only option is replacement. When upper tiers fail, however, application stakeholders must choose between deploying new devices that leverage different supporting infrastructure (obsoleting otherwise functional devices) or deploying replacement gateways or backhaul to resurrect extant devices. In some cases, such as the sunset of 2G wireless technologies, device owners have no option: a fixed resource (spectrum) that they do not own or control is taken away, and devices must be replaced.

One way to guarantee that a device will not be retired prior to its natural, functional obsolescence is vertical integration—own and operate all the supporting infrastructure. Indeed, consider the Voyager I probe, which is by far the longest existing, continuously operating, untouched device. It was launched in 1977 and can still communicate with NASA's Deep Space Network [25]. The extreme cost (in both time and money) required to replace Voyager justifies continued investment in its supporting infrastructure.

Building infrastructure can be costly, and for many applications supporting gateways and especially backhaul may already exist. Indeed, for smaller-scale stakeholders, vertical integration is likely infeasible. For larger stakeholders, such as municipalities, the economies of scale may quickly become accessible. Planners should consider the amortized cost of shared infrastructure over the cost of many applications. Systems should follow the architectural guidelines of the prior sections and use only non-vendored, standards-compliant gateways and backhauls. This allows early deployments to bootstrap off of extant infrastructure, while later permitting cost reductions with a transition to self-owned infrastructure. As the number of deployed devices grows, so does the cost of replacing them, and as a result the implicit lock-in to their supporting infrastructure model. For larger entities then, there will always be a tipping point where the cost of deploying vertically owned and managed infrastructure is lower than the cost of replacing devices. As this cost will vary with time and technology, it is imperative that all entities retain the option of self-reliance, should that become the cost-effective model for their application needs.

Takeaway: Stakeholders who deploy century-scale devices, and are responsible for managing their applications, should reserve the option of vertical integration, which is enabled by runtime-swappable gateways and backhaul.

### 4 A 50-YEAR EXPERIMENT

As a test of some of the principles discussed in this work, we commence a (hopefully very) long-lived experiment. The experiment consists of edge devices designed to last "forever,” a universal infrastructure guided by our principles for communication, and a backend server which publishes a web page with the data collected from the devices.

The top-level constraint for the experiment is that once the edge devices are deployed they are never touched again. This is representative of the infeasibility of maintenance on vast numbers of edge devices at scale. Similarly, we hope to minimize (but expect) intervention in the gateway and backhaul infrastructure to maintain the operation of the system. Our initial metric for end-to-end uptime requires that some data arrives at some interval of time up to once a week that is publicly accessible at centurysensors.com.
4.1 Devices
We begin this experiment with only a modest number of simple devices. Over time, we imagine the steady addition of new instances and types of devices. This ease of device deployment is directly the benefit proposed by the availability of trusted, stable infrastructure.

For the initial experiment, we focus on energy harvesting, transmit-only sensors. These are devices with minimal security risk, as they are incapable of receiving data, but also of limited longitudinal trust, as their security and signing techniques can never be modified. As these are energy-constrained devices, they will use low power radio protocols—802.15.4 and LoRa to start—, which will require gateway support to forward data to the internet at large.

4.2 Gateways
For gateways, we use our two radio technologies to explore two design points. In the first case, we consider an “owned infrastructure” scenario, where we will deploy, maintain, and operate 802.15.4 gateways. For the second case, we consider a (hedged) version of a “third-party infrastructure” scenario, where we rely on extant, environmental LoRa gateways who pay to ferry sensor data. Here, we will leverage the emergent Helium network [16]. The advantage (and risk) of Helium is that it is a semi-federated network, which enables us to own and operate gateway devices that we could use to supplant infrastructure if the commercial network were to become unusable. In the initial deployment at least, the third-party option achieves its expected easier deployment, as we must deploy nothing more than the edge device.

While we aspire to set-and-forget gateways, this part of the experiment will allow for maintenance and upkeep.

4.3 Backhaul
Here again, we begin with two designs. For the owned 802.15.4 gateways, we are obliged to provide backhaul connectivity. This first deployment then mimics a “municipal-provided” backhaul, where we are able to leverage our campus network to provide robust, reliable, and free connectivity for our gateways. In the Helium case, the backhaul is largely opaque so long as third-party gateways remain operational. Preliminary experiments probing the Helium network find that Comcast, Spectrum, and Verizon are the ISPs for roughly half of the 12,400 gateways with public IP addresses.5 If we deploy our own Helium gateways in the future, this may present an opportunity to explore “commercial-provided” backhaul as well.

4.4 Management
The end-to-end system will require maintenance before the fifty year mark. Our experiment stipulates that devices remain untouched, but if they do fail, we will document, diagnose, and replace them—this is intended as a living study.

For the 802.15.4 gateways, we rely on the reliability of a (networked!) Raspberry Pi-class device.6 The initial application supported by the gateway is transmit-only, which would allow it to be aggressively firewalled and limit the security risk of not attending to updates. Unidirectional gateways limit the utility of our deployed infrastructure, however. Thus we anticipate a more traditional server model, with the requisite upkeep of any public-facing, networked device.

For the Helium gateways, we are relying on the success of an emerging technology. We will, of course, also need to pay for this service. One interesting property is that the price of data once purchased is fixed. For one device to send one (up to 24-byte) packet every one hour for 50 years will cost 438,000 data credits. We can provision a dedicated wallet today with a conservative 500,000 data credits for just $5 USD. In theory then, we can provision a device and prepay its data to enable unattended operation for 50 years.

Finally, there is the data endpoint itself. We have intentionally focused less on this aspect, as long-lived cloud services are comparatively well-understood. Still, we will have to establish and maintain a reliable endpoint for data collection as well as potential data retention and resiliency.

4.5 Expected Outcomes
We intend to use the data display webpage as a living, public experimental diary. Here we will document any maintenance or changes we have to make to devices, gateways, or backhaul infrastructure to sustain operation. This includes recurring costs and periodic, predictable efforts that go into sustaining this system (e.g. one certain event: the maximum domain lease is 10 years [18]). The hope is that it can serve as a guide for real-world maintenance challenges of long-lived systems. It will also include a log of the experimenters, as the nature of a 50-year experiment is such that those who start it will most likely be retired by the time it is complete!

5. FINAL THOUGHTS
This paper does not answer the question of how to build century-scale systems. The management of long-lived, traditional infrastructure is a challenge that society is still dealing with today. As we introduce electronics and intelligent systems to infrastructure, the onus is on us as technologists to ensure that we make infrastructure management and maintenance better, and not worse.

---

550% of nodes belong to just ten ASes, but the long tail extends to nearly 200 unique ASes providing connectivity. More detailed analysis of the Helium backhaul is left to future work.

6In at least one case, a non-networked Raspberry Pi has operated unattended for nearly eight years and counting [5].
REFERENCES


