MantisMail: Green Content Delivery for Email

Nan Deng, Chen Li, and Christopher Stewart
Computer Science and Engineering
The Ohio State University

Xiaorui Wang
Electrical and Computer Engineering
The Ohio State University

Fiber optics have narrowed the gap between an ISP's total web traffic and bandwidth provided by high capacity links. Now, datacenters connected to ISPs via fiber links can quickly deliver web content to the ISP's users without placing edge servers within the ISP. Due to economies of scale, these datacenters can host web content efficiently, making it cheaper to set up a content delivery network (CDN) that caches content to reduce delivery times. Increasingly, CDNs must do more than reduce response times to attract customers; They also have to keep prices low and invest in new features that add value to their service.

We propose a new value-added feature for CDNs: Carbon offsetting. Carbon offsets represent a unit of work that voids 1g of carbon emission. By buying carbon offsets, CDNs can void emissions caused by traffic that they intercept, including emissions spent getting content from source servers. These *green CDNs* would add value because the growing carbon footprint of the web is concerning. If datacenters worldwide were treated as a country, they would rank among the top 30 carbon emitters. Routing, storing, and serving email alone causes more than $1M$ metric tonnes of carbon emission each year [1, 2], even though email is less than 1% of Internet traffic.

Research Challenge 1: To offset the emissions caused by retrieving content from source servers, green CDNs must estimate the carbon needs of servers and network links they do not control. Prior studies have estimated carbon needs for specific workloads, but most of these studies confound the contributions of applications, hardware, and environment. Our approach captures each independently, allowing us to predict the effects of Moore's Law and growing renewable energy penetration.

Research Challenge 2: The simplest design for a green CDN is 1) cache content on a cluster of servers, 2) size the cluster based on hit rate and bandwidth costs, 3) estimate emissions, and 4) buy offsets as needed. But carbon offsetting increases costs, affecting the best cluster size for a green CDN. Further, these costs vary over time with energy prices and renewable energy production.

$$\text{Min: } \lambda P_{c02e}[E_{source} \times M(N)] + \lambda P_B + P_{srvr} \times N \quad (1)$$

$$\text{Subject to: } \mu N \geq \lambda \quad (2)$$

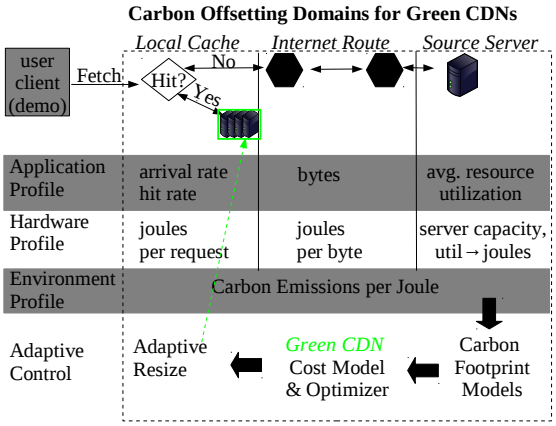

Figure 1: Data flow, profiling, and modeling in MantisMail—a green CDN. *At NSDI, we will demo MantisMail on our email accounts and allow attendees to test their own.*

Equations 1 & 2 preview our cost model that helps green CDN managers size their local cluster. N is the CDN cluster size. λ is the arrival rate of user requests. μ is the processing rate per instance; $M(N)$ is the miss rate function. E_{source} is the carbon footprint for getting content from the source. P_{c02e},$P_B$andP_{srvr} are the price of carbon offsets, outgoing bandwidth, and CDN servers. Our goal is to minimize cost with low response times.

MantisMail: Figure 1 depicts data flow, adaptive control, and carbon footprint modeling in *MantisMail*, our proof-of-concept green CDN. End users sign up for MantisMail by registering their email ID and source IMAP server address. Within their email clients, users configure MantisMail as their IMAP server. MantisMail caches entire IMAP requests using the Redis key value store, serving cached content when possible. Requests that can't be cached are sent to the user's source IMAP server. Note, MantisMail supports SSL and does not store passwords locally. Cached emails are encrypted.

Recently, the authors began accessing their email through MantisMail, offsetting 295g C02e per day (about 1 mile in an econ. car). While email is personal, cache hits occurred for users with multiple devices (12% hits). A content based cache could increase hits.

References

[1] ICF International. The carbon footprint of spam report, 2009.

[2] S. Radicati. Email statistics report:executive summary, 2010.