



Green Bytes: Sustainable Approaches to Digital Stewardship, A Practitioner's Perspective





There Are Limits...

Existing “Green” Preservation Practices

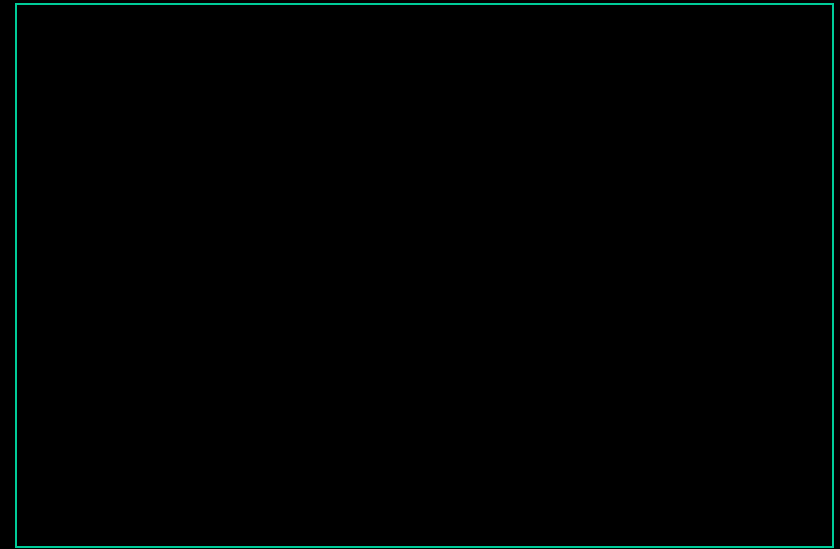
Tape Archives

“Virtual stacks” in...

Berms

Mountain vaults

- Leased salt mines
and stone quarries



storage
access

lower
higher

Leading “Green” Commercial Practices

Innovative physical infrastructure/cooling
Geographic distribution of facilities
High/er storage density
High/er temps

GreenGrid Key Metrics:

PUE, DCcE, ScE...

www.thegreengrid.org



Case Study: Internet Archive

www.archive.org



Challenges to Sustainability: OpEx

Initial Goal: Lower operational expenses by reducing power consumption

Base Metrics:

- Power Usage Efficiency
PUE
-



Power Usage Efficiency

$$PUE = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$

- Industry average for typical data center in 2008: PUE 1.5 – 2
 - Meaning: 100W of computing takes 50-100W of cooling

The “Old” Days (2004 - 2010)

- Traditional, air conditioned, data center facilities
- Rented - NOT easy to modify
- No separation between hot & cold air
- PUE (2010):
est. ~1.8
- 2004: 117 KW/PB
- 2010: 39 KW/PB
not including cooling



Container Experiment (2008-2012)

- Density made cooling more important/complicated
- Needed cold water, not air
- PUE: varies (est. ~2)
- 67 KW/PB
not including cooling

Cheaper and faster to
deploy than a traditional
data center

The Latest Chapter (2010 – present)

Reduce power consumption even further/increase server efficiency

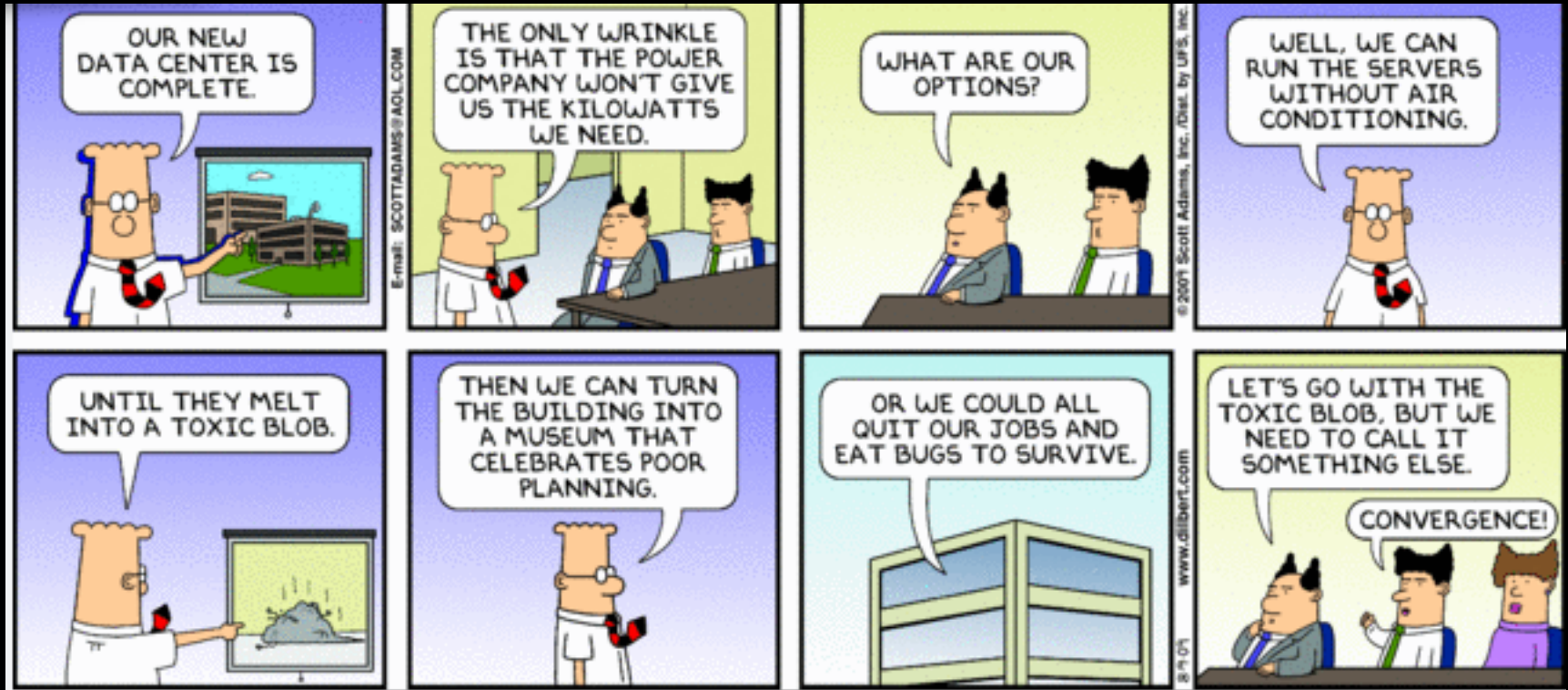
Racks should co-exist with humans, don't hide them...

Do not ruin buildings with ducts, A/C plant (in SF, 2 miles from the ocean w/49 wks per yr of natural cooling...)



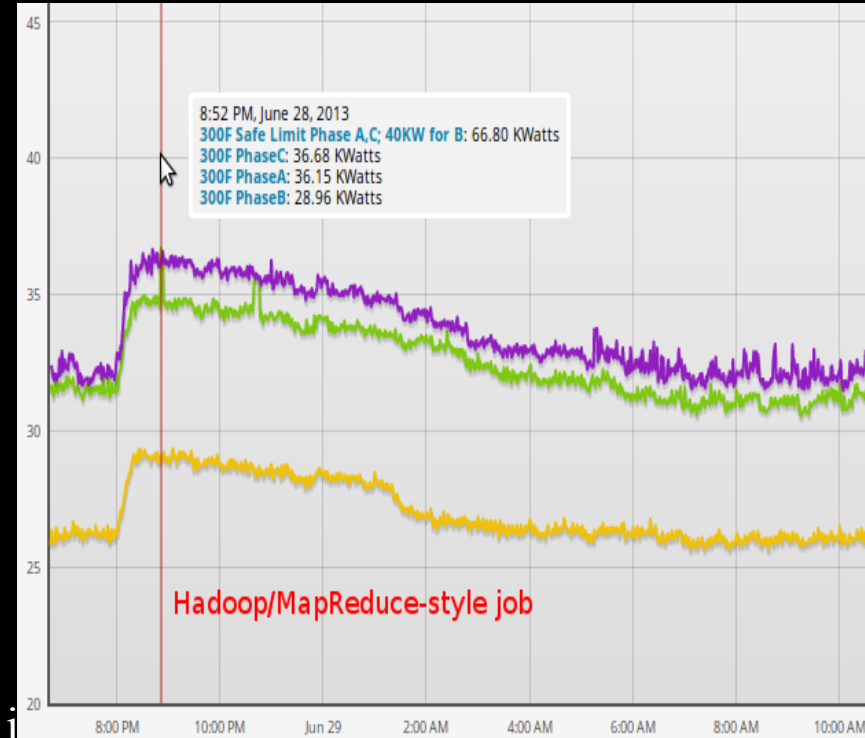
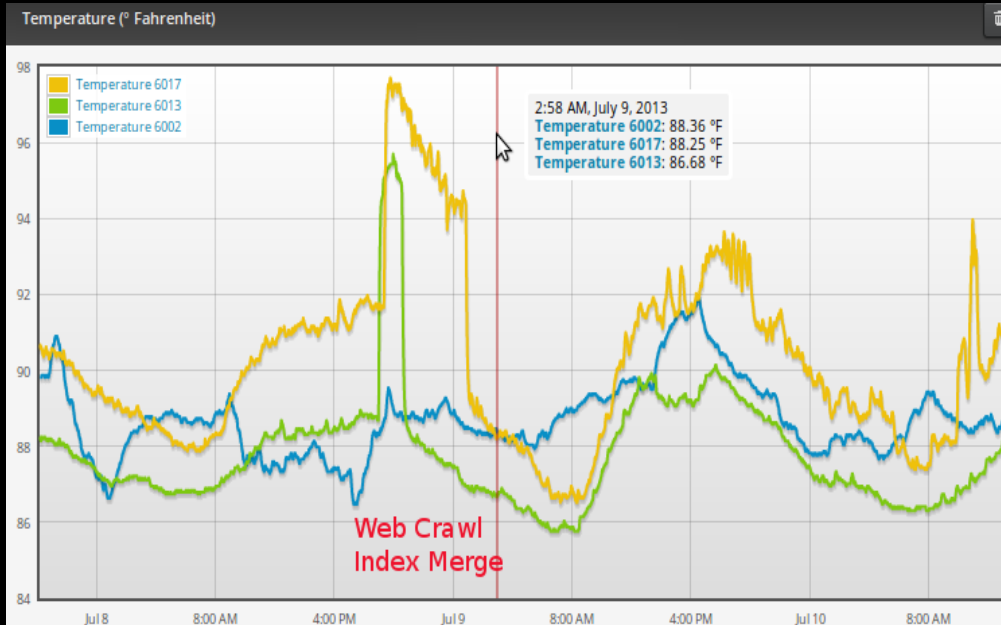
Make it as “pretty” as...
the Jedi Library from Star
Wars

Challenges Along the Way



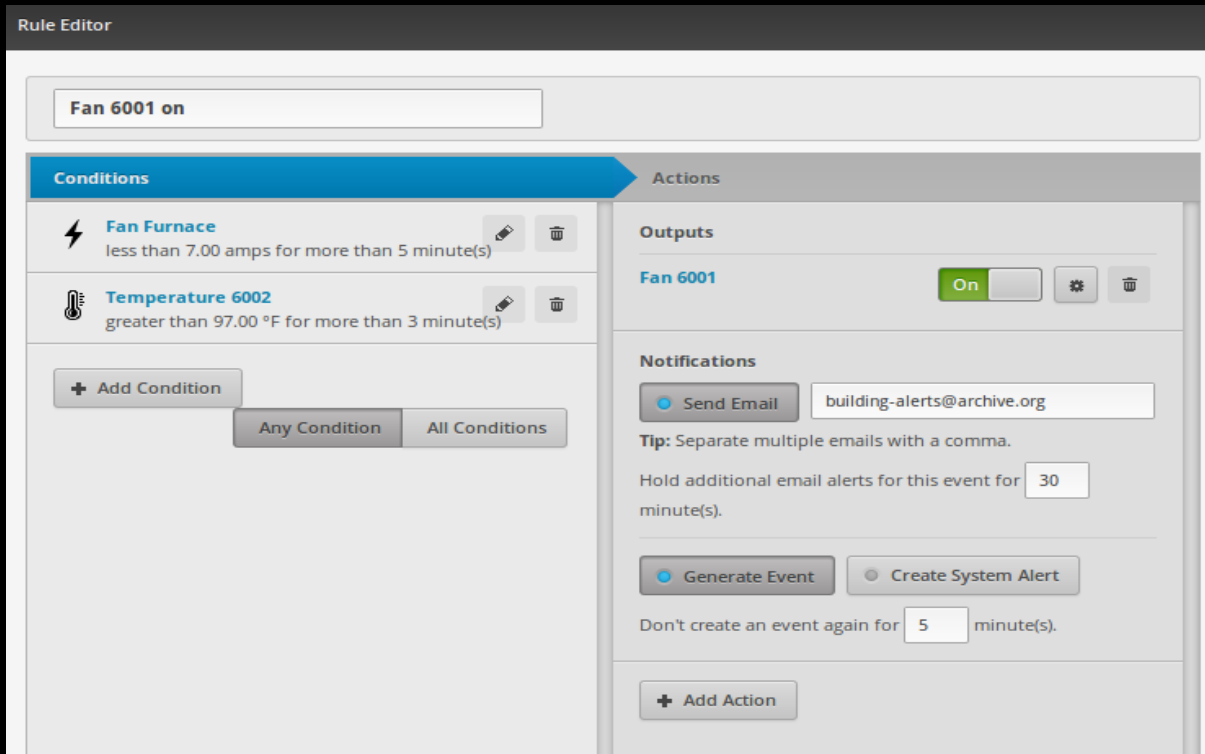
No support from the power company and no \$\$ to “buy” solutions to measure usage...So, we built a whole-data-center, real-time, networked power meter for < \$500...

Ongoing Engineering: Data Driven Optimization



- Measure
- A/B tests
- Correlate power, temperatures, compute jobs
- Shift big compute jobs to avoid noon heat (and peak utility pricing)

Ongoing Engineering: Software + Networked Devices



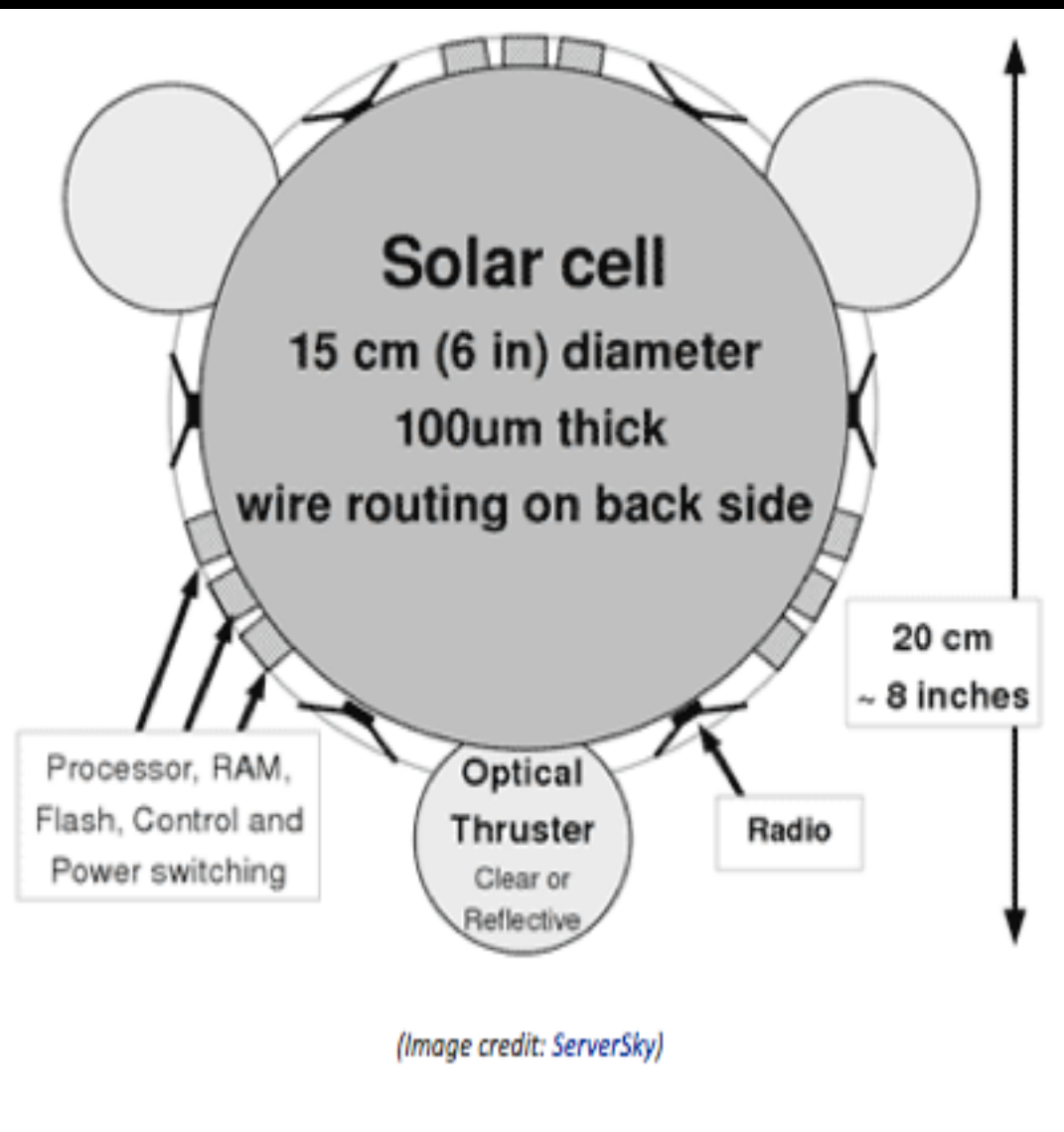
Shift from “dumb”
hardware to networked
devices plus software...

```
if (temperature > 97F)    # it's hot
  or (current(fan1) < 7A) # fan1 has failed
  then power_up(fan2)
  and alert(humans)
```

For Now...Success!

- PUE < 1 (heat gets reused)
 - 2010: 8.5 KW/PB
 - 2013: 2.8 KW/PB
(+1 KW, if running compute VMs)
 - Quiet(ish)
- Most photographed data center in the world ...





(Image credit: ServerSky)



500,000+ Books
500,000+ Moving Images
1,000,000+ Audio
2,000,000+ Recordings
3,600,000+ Hours of TV
350,000,000,000 news
+ eBooks/eTexts