

Towards a science and technology for sustainable cyberinfrastructure

Fernando J. Pineda

Dept. of Molecular Microbiology & Immunology, Dept. of Biostatistics
Johns Hopkins Bloomberg School of Public Health
Director: Joint High Performance Scientific Computing Exchange

Fernando.pineda@jhu.edu

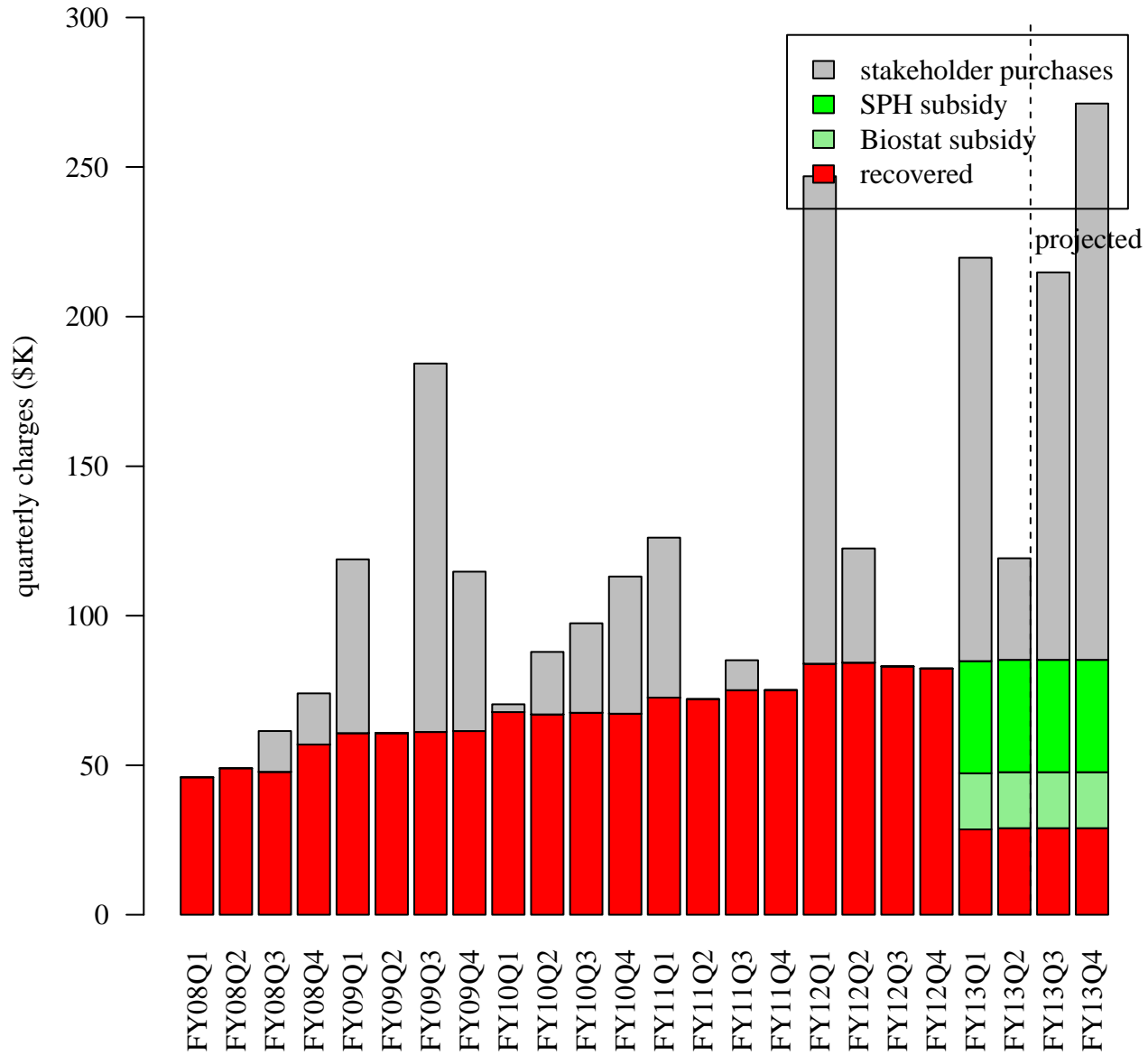
Sustainable Business Models for Sustainable Cyberinfrastructure Workshop II,
Oct. 2 2013

whomai

- Assoc. Prof., Dept. of Molecular Microbiology & Immunology
 - Dept. of Biostatistics
 - Division of Health Science Informatics
- Theoretical Physicist, Supercomputer training (Thanks NSF/DoD/Boeing!)
- 10 years ago I used my startup funds to hire a sysadmin and to buy a small cluster (\$30K), then expanded into a vacuum...
- Director: Joint High Performance Computing Exchange
 - Serves: School of Public Health, School of Medicine & Lieber Institute
 - 450 user accounts (250 unique users/mo).
 - 1000 cores/10TB ram in production
 - 1000 cores/10TB sitting in boxes in the server room (don't ask)
 - ~1PB (formatted) disk storage (includes 700TB DIY ZFS storage system)
 - Backup to IBM Tivoli tape system ('sold' requests for ~ 400TB of backup)
 - 1.85 FTEs, ~\$400K/yr operations, ~ \$1.5M worth of hardware
- “Postmodern” condo model, 100% chargeback until July 2012

Performance of “postmodern” condo model

100% chargeback until July 2012



Assumption

- Individual PIs know what they need far better than any committee purporting to represent them.
- PIs can include the cost of computing in their grants and this is business as usual in NIH funded institutions
- Apparently including computing charges is problematic for NSF funded investigators -- but it shouldn't be!

View Cyberinfrastructure facilities as Common Pool Resources

- Common Pool Resources?
 - A concept usually associated with natural resources, e.g. fisheries, but also man-made systems, e.g. irrigation systems, but the notion fits Academic HPC facilities too!
 - Many of our headaches are common to CPR
 - Garrett Hardin 1968 -- Tragedy of the Commons
 - The most widely cited paper in Science Magazine?
 - Elinor Ostrom -- 2009 Nobel Prize in economics
 - For work on CPRs. She and collaborators identified design principles that are prerequisites for a sustainable CPR

Design principles for sustainable CPRs (with apologies to Ostrom et al.)

1. **Clearly defined boundaries** – What are the boundaries of the CPRs and who are the members of the individual consuming entities?
2. Congruence between consumption/provision rules and local circumstances
3. **Collective-choice arrangements** allowing for the participation of most of the consumers in decision making processes
4. **Effective monitoring** by monitors who are part of or accountable to the consumers. Who monitors the monitors?
5. **Graduated sanctions** for consumers who do not respect community rules
6. **Conflict-resolution mechanisms** that easily resolve conflicts between consumers as well as between consumers and governance organizations.
7. **Consumers must be allowed to self-organize** without interference from governance
8. **Hierarchical Systems** -- provisioning and consumption must allow for multiple layers of nested CPRs, with small, local CPRs at their bases.

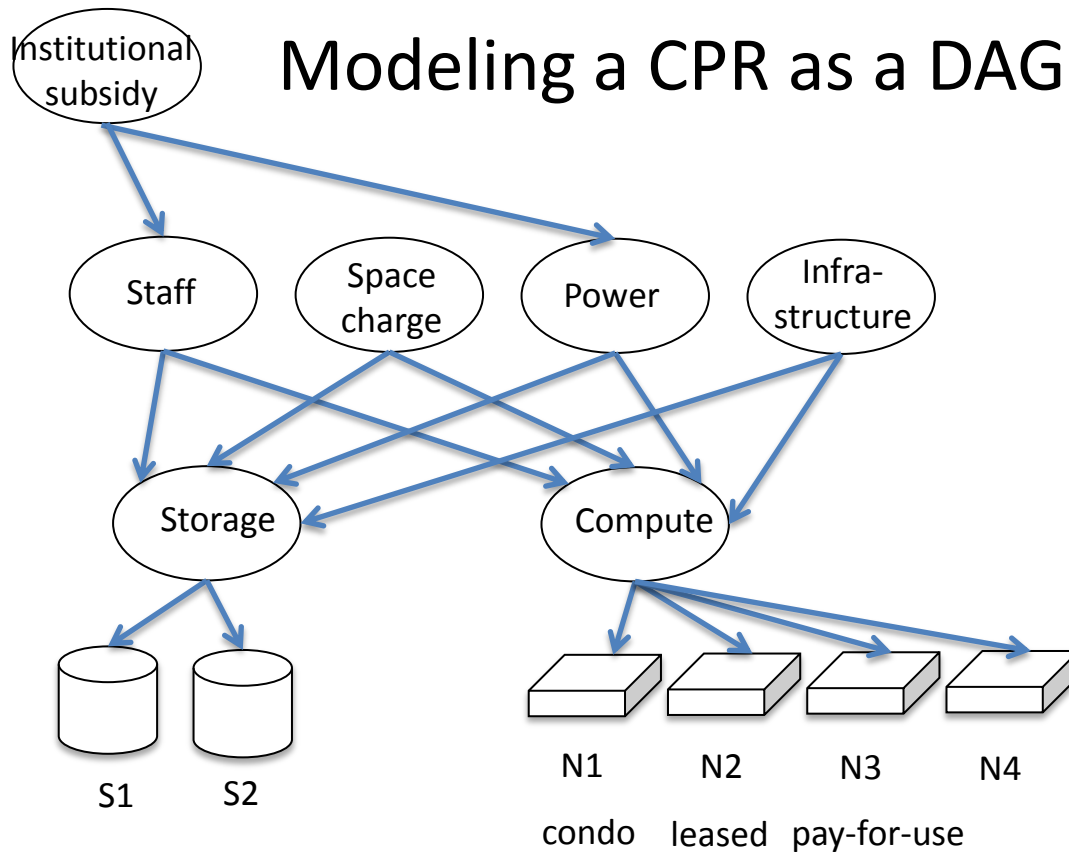
(see e.g. <http://www-personal.umich.edu/~rdeyoung/tragedy.html>)

Principles are nice, but they are useless if they can't be implemented

- Fortunately Cyberinfrastructure CPRs are:
 - highly structured
 - Easily instrumented/monitored
 - Data rich
- You could not ask for a better laboratory for developing tools for sustainable management of common pool resources

Implementation

- Partition the problem into three essentially independent tasks
- CPR modeling
 - Sub CPRs, e.g. CPU, GPU, storage, tape backup, etc.
 - Cost accounting in space and time as provisioning changes
- Allocate costs in proportion to *relative* consumption
 - Measure consumption
 - Everyone who uses a resource is charged equally for consumption. Affiliation is irrelevant.
 - Generate “bills” even if individual faculty don’t see them because someone has pay
 - Stakeholders own resources, but need not be users or individual faculty
 - Surprise: A simple rule allows for constant charge and charge-by-cycle to coexist in the same system (one size does not fit all).
- Policy
 - Policies are used to govern who gains access to which parts of the system and when.
 - Policies are used to determine rules for charge outs



Every vertex has costs (positive for expenses, negative for subsidies)

Costs and subsidies can be injected anywhere

Costs propagate downwards along the edges to the leaves.

The leaves correspond to resources that can be consumed

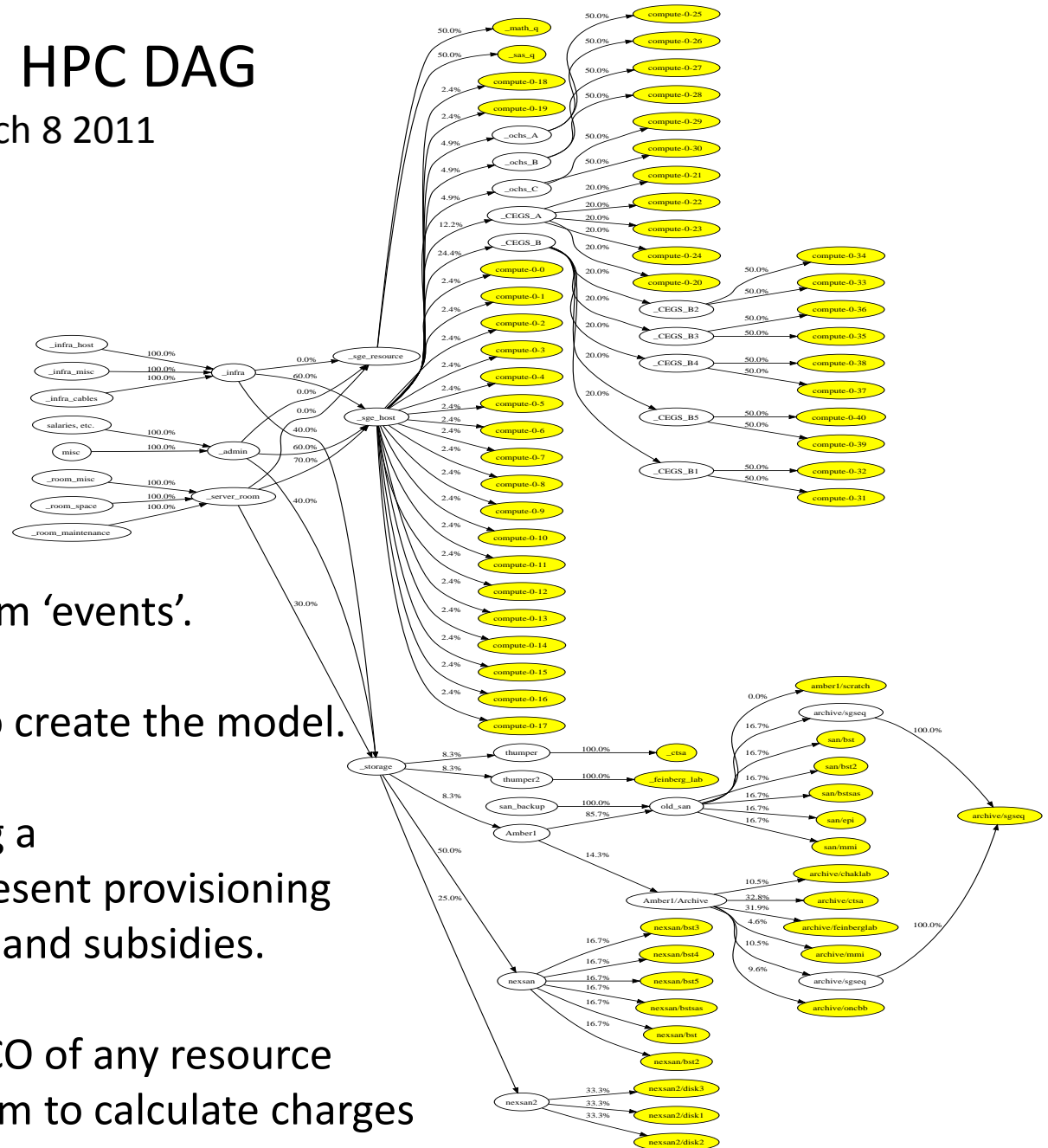
Consumers are allowed to own individual leaves to satisfy their research agendas

The model must be easy to maintain otherwise sysAdmins won't keep it up-to-date

The model can be published for all to see – no secrets!

A real-world HPC DAG

HPSCC, March 8 2011



The HPCDAG is compiled from 'events'.

Given data, takes < 1 hour to create the model.

Updated quarterly by adding a handful of text lines to represent provisioning changes as well as expenses and subsidies.

Integrate over time to get TCO of any resource
Incorporate into billing system to calculate charges

Example policy for charging

- Each resource (or set of resources) is treated as it's own sub-CPR which is owned by a stakeholder.
- Stakeholders can be individual faculty, departments or institutions
- The i-th user is charged in proportion to their relative usage of the CPR.

$$c_{i,r} = C_r \frac{u_i}{\sum u_i}$$

- Stakeholders who do not share their resource see a fixed predictable charge for using their resource.
- Stakeholders who share their resource receive a discount
- Non-stakeholders are charged (essentially) in proportion to cycles.
- Stakeholders can buy 'surge' capacity from other stakeholders as needed
- Users/stakeholders allowed to self-organize **locally** to package charges into whatever payment model works for them, free, pay-for use, etc.

Parting shots

- Tools under development with proof of concept in real-world setting.
- There is an interesting research/engineering problem here.
- The HPCCPR problem is probably at least as interesting as the job scheduling problem. Why hasn't anyone tackled it?
- Can I get NSF or NIH funding to pursue this?
- Oh right...no one is answering the phone.