Decreasing Data Center Energy Use through Network and Infrastructure Control

Real-Time Optimal Control

PI: Ron Moon (703) 488-5044 Ron.Moon@fts-intl.com

Project Name: Decreasing Data Center Energy Use through Network and Infrastructure Control

- **Lead organization:** BAE Systems
- **PI:** Ron Moon
- **Participants:**
  - Cisco Systems (vendor)
    - Data Center & VoIP research experts
  - TRG Consulting (vendor)
    - Real-Time Optimal Control & math modeling experts
- **Project:**
  - Start: March 31, 2010
  - Completion: March 22, 2011
- **Project type:** Concept definition

Transformation / Game Changing

BAE SYSTEMS

The bounds continues to expand due to:

1. Advances in mathematics
2. Moore’s law & increased hardware control features

Heterogeneous, complex system control is now possible

However, it is very difficult to get people excited about math!

Project Background

BAE SYSTEMS

"Widespread underutilization of servers is one of the most often-cited reasons for suboptimal energy efficiency in data centers."

- EPA Report to Congress on Server and Data Center Energy Efficiency

- Servers typically operate at an average processor utilization level of only 5 to 15%
  - Consuming 60 to 90 percent of its maximum power
- Existing Routing and Telecom Data Center power control systems are implemented in a disjointed manner
  - Range from manual to automated hierarchical systems
  - Existing automated control systems:
    - are often crude, time-based implementations
    - operate at each hierarchy layer attempting to minimize equipment power consumption but often adversely impacting the attempts of adjacent power control layers, resulting in reduced power efficiency

The hardware and software capability is emerging to implement automated systems-of-systems power management control at the macroscopic level

State of Practice

Supply, supply, supply…"**Justin**" Principle

- **<u>Just-in case</u>** the center receives a surge in processing requirements
  - We can't violate our QoS metrics, we need to be ready!
- Existing, disjointed control systems cannot respond fast enough to maintain QoS

Transform from a dumb supply approach to…

Energy Savings: State of Art

BAE SYSTEMS

…an autonomous, intelligent Real-Time Optimal Control

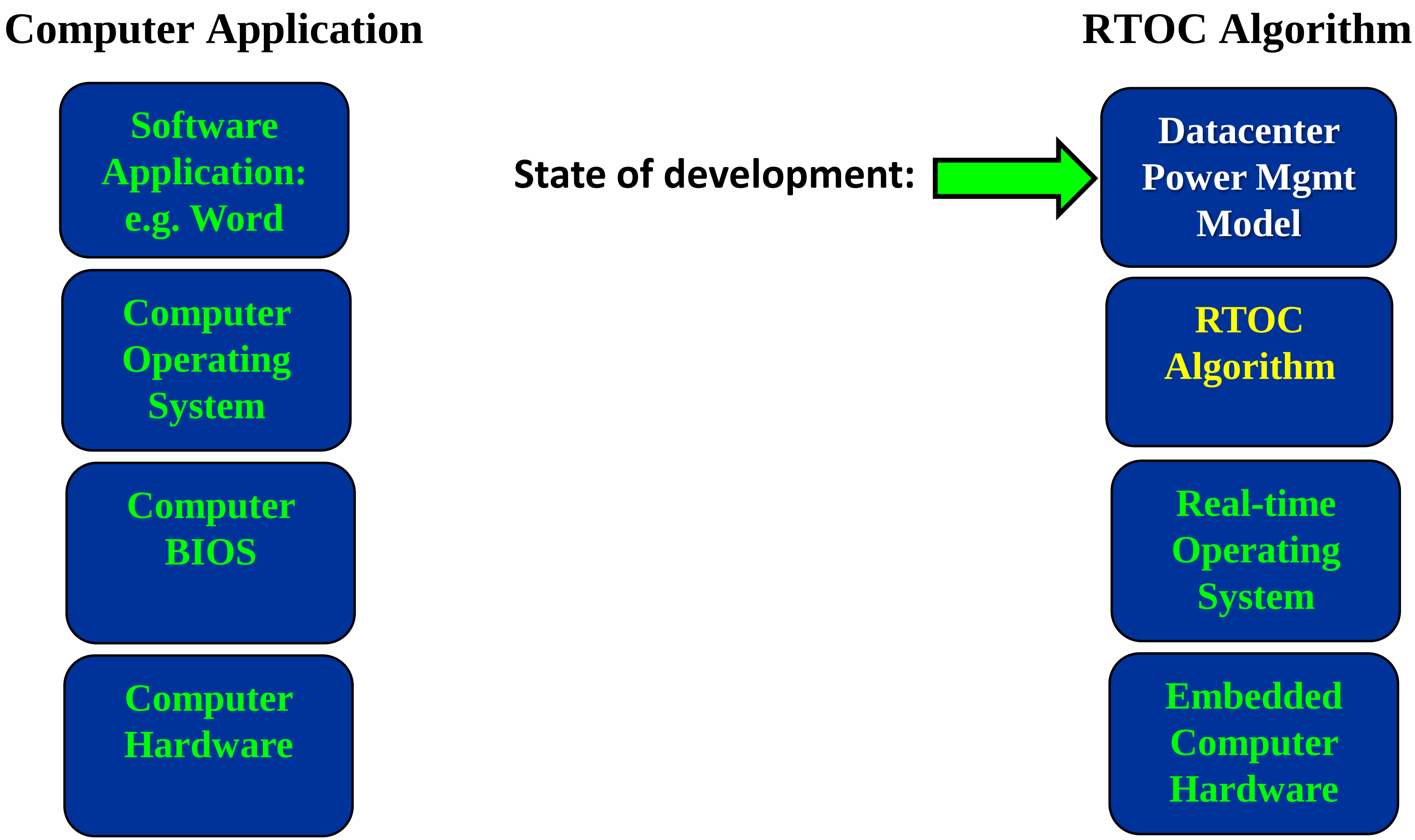
Power down excess capacity HAVC units

Power down excess servers, fans and support devices

What Are Optimal Control Problems?

Fundamental mathematical problems that have wide applications in almost all branches of engineering: Aerospace, Chemical, Petroleum, Power, etc.

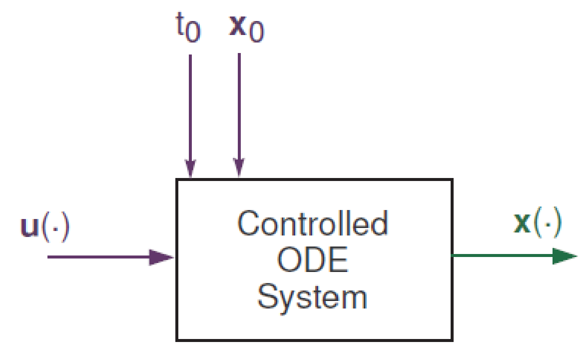
Algorithm Implementation Analogy



Approach: Problem Formulation

E: Event Cost Function (Mayer cost)

F: Running Cost Function (Integrand of the Lagrange cost)



$$\mathbf{x} \in \mathbb{X} = \mathbb{R}^{N_x}, \quad \mathbf{u} \in \mathbb{U} \subseteq \mathbb{R}^{N_u}$$

$$\begin{cases} \text{Minimize} & J[\mathbf{x}(\cdot), \mathbf{u}(\cdot), t_f] = E(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} F(\mathbf{x}(t), \mathbf{u}(t), t)\, dt \\ \text{Subject to} & \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \\ & \mathbf{x}(t_0) = \mathbf{x}^0 \\ & t_0 = t^0 \\ & \mathbf{e}(\mathbf{x}_f, t_f) = \mathbf{0} \end{cases}$$

Find u(t) such that "*J* " (cost) is minimized

Constrained Dynamic Optimization…I was told there would be no math!

Project Objective / Goal

BAE SYSTEMS

Concept definition study

- Explore and document the technical requirements and impediments to implementing a RTOC algorithm in Routing and Telecom Data Centers
- Identify the critical elements of the RTOC algorithm requiring feasibility demonstration
- **Translate Routing and Telecom Data Center behavior into preliminary technical specifications and model**
- Increase understanding and mitigate market introduction and technical risks
- Enhance community understanding regarding the potential of a RTOC algorithm to reduce Routing and Telecom Data Centers' energy consumption
- Document research results
- Develop R&D plan for Gate 2 research approval to enter Stage 3 (Concept Development)

Common error: rushing through model development

Results & Status: Data Center Math Model

- A data center processes a given number of apps A(t)
- All apps are processed in blades
- A given blade is classified by its location, functionality, and customer

***Location Classification*:**

- The location of a blade is identified by a triplet, (rack (**r**), chassis (**c**), blade (**b**))
  - (2, 4, 6) identifies blade # 6 on chassis # 4 in rack # 2

Functionality Classification:

- Identifying and mapping blade types to location triplet (r, c, b):
  - I_p, I_n, I_s: set of all processing, network and storage blade triplets, respectively
- Thus, any blade can now be located by a triplet and identified (by association) to be processing, networking or storage

Customer Classification:

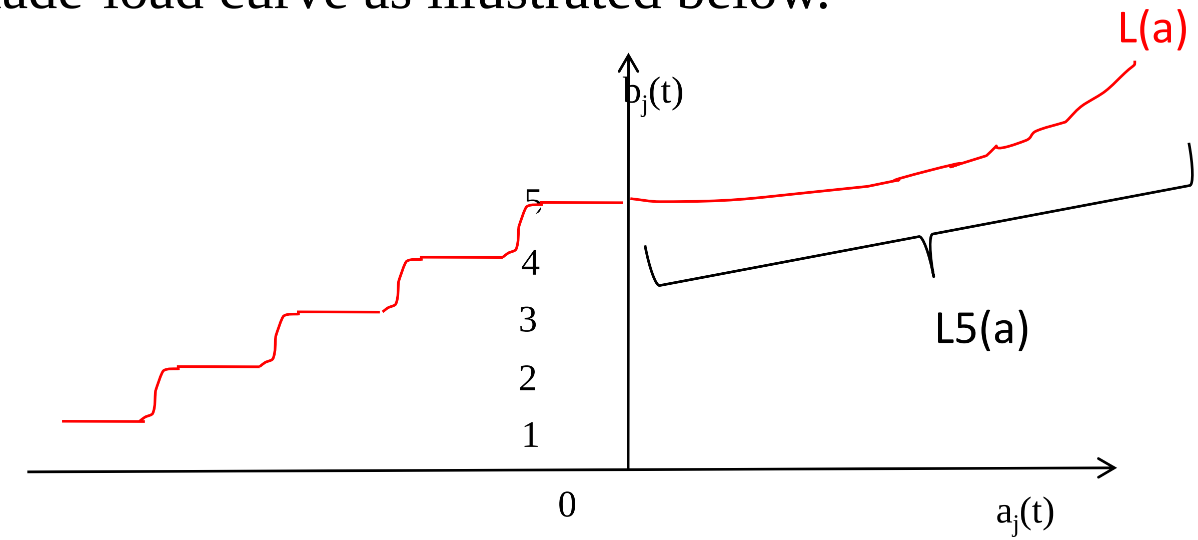
- Blades can also be classified in terms of customers to which its assigned
- Let J_k be the set of all triplets (r, c, b) that identify a blade to serve customer k
  - Apps from customer k_1 will not be allocated to any blades that belongs to customer k_2, where $k_1 \neq k_2$

Blade identification scheme accounts for Data Center construct, functionality and customer funded QoS

Results & Status: Generalized Blade Load Concept

BAE SYSTEMS

- In order to perform unified mathematical operations, we propose the concept of negative apps to form a generalized blade-load curve as illustrated below.



- If negative apps are running on a blade, it takes on the discrete states, P0-P4.

The concept of negative apps extends the P5 notion and allows using apps to autonomously control the power state of the blade

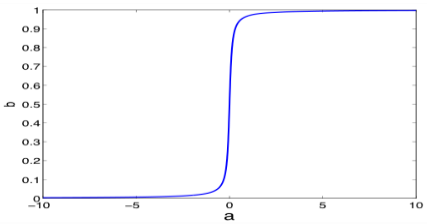
Results & Status: Generalized Blade Load Model

BAE SYSTEMS

- The blade-load curve b=L(a) can be made smooth so that the partial derivative of L(a) with respect to a(t) exists

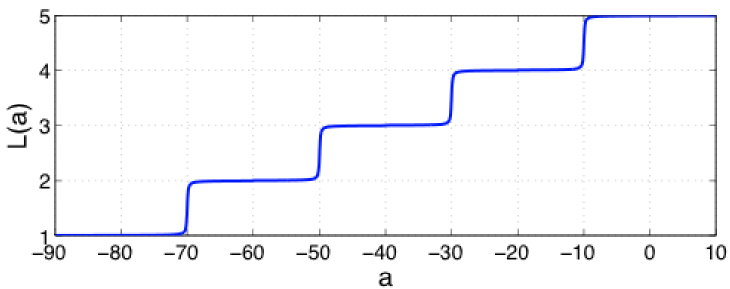
- We will use the following function as a building block to smooth the discrete part of the states

$$b = L(a) = \frac{1}{\pi}\tan^{-1}(c_1 a) + 0.5$$



Constant C_1 is used to control how smooth the function is. When C_1 = 10, the shape of the smoother function is shown in this figure.

$$L(a) = \begin{cases} \frac{1}{\pi}\tan^{-1}(10(a+70)) + 1.5, & \text{if} -90 \le a \le -60 \\ \frac{1}{\pi}\tan^{-1}(10(a+50)) + 2.5, & \text{if} -60 \le a \le -40 \\ \frac{1}{\pi}\tan^{-1}(10(a+30)) + 3.5, & \text{if} -40 \le a \le -20 \\ \frac{1}{\pi}\tan^{-1}(10a+10) + 4.5, & \text{if} -20 \le a \le 10 \end{cases}$$



- The blade is now controlled by the apps, real and negative (fake), running on it from the blade dynamics equation, $\dot{b}_j = \partial_a L(a_j)\dot{a}_j$

Controlling the apps running on the blade, controls the power state of the blade

Results & Status: Model of Incoming Apps

BAE SYSTEMS

- A data center accepts a given number of apps, A(t), at time t
- The allocation of A(t) to the blades is a decision variable
  - This is currently accomplished by the Load Balancer
- A(t) belongs to K number of customers. Let $[A(t)]_k$, k=1,2,…,K be the apps form customer k. Then

$$A(t) \;=\; [A(t)]_1 + [A(t)]_2 + \cdots + [A(t)]_K$$

- $[A(t)]_k$ is a vector of “processing”, “network” and “storage”. Let $[A_p]_k$, $[A_n]_k$ and $[A_s]_k$ be the number of incoming processing apps, network apps and storage apps from customer k; then,

$$[A(t)]_k \;=\; [A_p]_k + [A_n]_k + [A_s]_k$$

$$\begin{aligned} A(t) &= [A(t)]_1 + [A(t)]_2 + \cdots + [A(t)]_K \\ &= A_p(t) + A_n(t) + A_s(t) \end{aligned}$$

Sum of apps from all customers

Results & Status: Blade Dynamics w/Incoming Apps

BAE SYSTEMS

- It can be shown that the apps dynamics is given by,

$$\dot{a}_j(t) = \begin{cases} k_p(a_{p,j}(t), b_j) + k_n(a_{n,j}(t), b_j) + k_s(a_{s,j}(t), b_j) \\ \qquad + a_{p,j}^{in}(t) + a_{n,j}^{in}(t) + a_{s,j}^{in}(t), & \text{if } a_j(t) \geq 0 \\ l(a_j(t), s), & \text{if } a_j(t) \leq 0 \end{cases}$$

- The blade dynamics is $\dot{b} = \partial_a L(a(t))\dot{a}$

- Substituting the apps dynamics in the blade dynamics, we have,

$$\begin{aligned} \dot{b}_j(t) &= \begin{cases} \partial_{a_j} L(a_j) \cdot [k_p(a_{p,j}(t), b_j) + k_n(a_{n,j}(t), b_j) + k_s(a_{s,j}(t), b_j) \\ \qquad + a_{p,j}^{in}(t) + a_{n,j}^{in}(t) + a_{s,j}^{in}(t)], \quad \text{if } a_j(t) \geq 0 \\ \partial_{a_j} L(a_j) \cdot [l(a_j(t), s)], \quad \text{if } a_j(t) \leq 0 \end{cases} \\ a_j &= a_{p,j} + a_{n,j} + a_{s,j} \\ j &= (1,1,1), \; \cdots, \; (N_r, N_c, N_b) \end{aligned}$$

Results & Status: Decision Matrices for sundries

BAE SYSTEMS

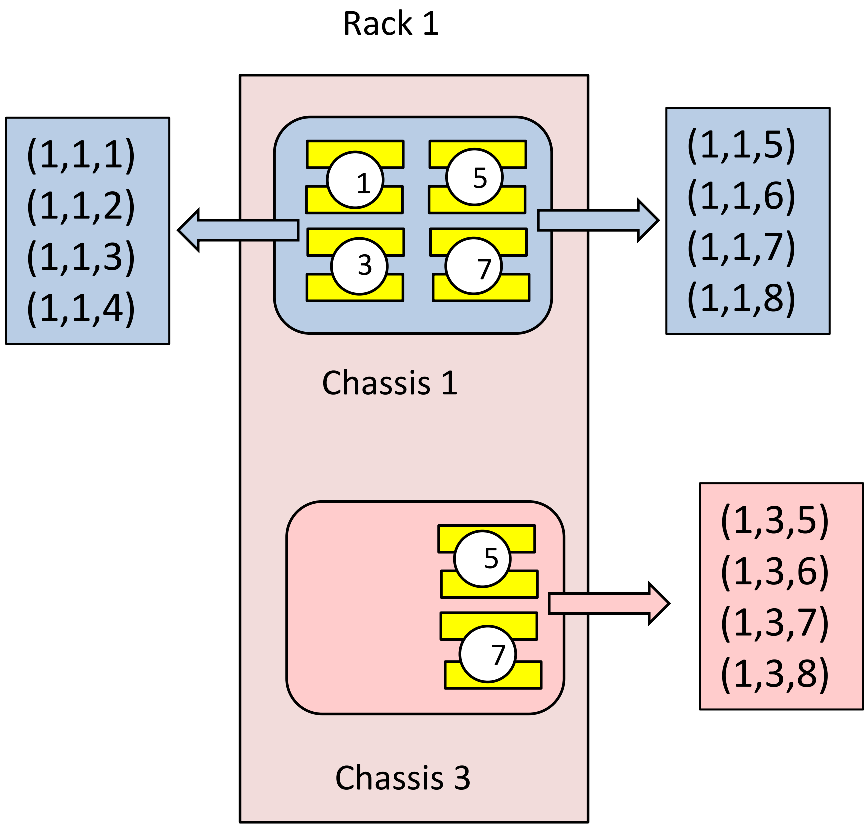
- The "dynamics" of the fan can be modeled via decision variables
- To perform this decision, a number scheme similar to the blades are adopted.

If $b_{(r, c, b)} + b_{(r, c, b+1)} = 0$, then $f_{(r, c, b)} = 0$ if (r, c, b) is an index set of the Chassis Fan Index

Thus $f_{(r, c, b)}$ (zero or nonzero) implies if the corresponding fan is turned off or on

A similar scheme can be constructed for the rack fans

The decision matrix is tailored to a specific data center design/configuration



Similar mapping and indexing scheme is applied to HVAC equipment

Value Proposition for End User

BAE SYSTEMS

- Cloud computing user
  - Minimize the total number of virtual machines running at any given time
  - Autonomously move your application to computing centers throughout the world with the lowest operating costs
    - Power, services, labor rates
  - Maintain customer QoS and control operating costs
- Minimize data center power consumption
  - Conversely, maximize data center productivity/earning capability

The basic framework of the control algorithm remains the same, it's the level of deployment and the cost function that changes

Estimates of Energy Savings

BAE SYSTEMS

- Based on 2006 Congressional information
  - Per Volume Server in Data Center – 985 kWhr/year - 10M in 2006
    - Technique applies to mid-range and high-end servers, but not included in calculations
  - Assumes 5% server growth, 60% market capture after 2013 introduction
  - HVAC savings additional
- Impact Projections Model – non-zero rows shown:

Impacts: New Technology vs. Current Technology

Real-Time Optimal Predictive Power Control of Routing and Telecom Data Centers

Impact By Year	2011	2015	2020	2025	2030	2035	2040
Cumulative Units installed	0	284,791	2,657,439	11,352,713	19,601,455	26,037,732	33,382,822
Annual New Units Installed	0	108,199	894,184	2,047,299	1,376,955	1,309,492	1,599,643
Annual Installations Total	0	108,199	894,184	2,047,299	1,376,955	1,309,492	1,599,643
Market penetration	0%	3%	21%	71%	96%	99%	100%
ANNUAL SAVINGS							
Energy Metrics							
Total primary energy displaced (trillion Btu)	0.00	2.32	20.63	88.15	152.19	202.17	259.20
Direct electricity displaced (billion kWh)	0.00	0.28	2.61	11.17	19.29	25.62	32.85
Financial Metrics							
Energy-cost savings ($MM/yr)	0.00	16.73	154.34	674.74	1174.11	1559.64	1999.60
Environmental Metrics							
CO Displaced (Metric tonnes)	0.00	46.76	429.80	1836.12	3170.22	4211.18	5399.13
Carbon Dioxide emissions displaced (MM TCE)	0.00	0.05	0.40	1.71	2.96	3.93	5.04
SO2 displaced (Metric tonnes)	0.00	656.30	5412.13	23120.88	39920.22	53028.31	67987.28
NOx displaced (Metric tonnes)	0.00	413.16	3553.40	15180.31	26210.14	34816.42	44637.93
Particulates displaced (Metric tonnes)	0.00	14.03	115.21	492.17	849.78	1128.81	1447.24
VOCs displaced (Metric tonnes)	0.00	5.02	46.29	197.76	341.45	453.57	581.52

Jobs / Employment

BAE SYSTEMS

- Project direct jobs/employment impact:
  - Efforts in this high technology field of research, development and deployment creates high-paying jobs, vice the low-paying job growth the economy is currently experiencing
- If this technology is commercially successful:
  - Reduced data processing energy usage translates directly into improved company Return on Assets (ROA)
    - Providing additional funds to reinvest in company growth
  - Spin-off algorithm applications for other business areas
    - Cloud computing
    - Cyber security
    - Medical treatment
    - Autonomous devices and systems (Robots, cars, sensor networks)
    - Factory process control
    - Hybrid vehicles

What's left?

BAE SYSTEMS

- Final Scientific/Technical Report
- Complete and submit required supplemental reporting
  - Final SF-425, Federal Financial Report
  - Patent Certification
  - Property Certification
  - Annual Indirect Cost Proposal – **N/A**
  - Annual Inventory of Federally Owned Property – **N/A**

The paperwork

Summary

BAE SYSTEMS

- Explored and documented the technical requirements and impediments to implementing a RTOC algorithm in Routing and Telecom Data Centers
- Identified the critical elements of the RTOC algorithm requiring feasibility demonstration
- Translated Routing and Telecom Data Center behavior into preliminary technical specifications and model
  - **Verified the data center can be modeled as a DAE**
- Increased understanding and mitigated market introduction and technical risks

Next step: determine/verify model curve responses through experimentation

After ITP-Sponsorship

BAE SYSTEMS

- BAE Systems is investigating opportunities to apply the control algorithm in its traditional DOD product line
  - Economic conditions will not support continued pursuit of the algorithm development in the DOE cost share development environment

Questions?

PI: Ron Moon (703) 488-5044 Ron.Moon@fts-intl.com