MIT Data Center Trade Off Analysis

Jeremy Kepner
MIT Lincoln Laboratory

December 2005

Outline

- **Introduction**
  - *Motivation*
  - *Study Goals*
  - *How did I get here*
  - *Summary of Options*
- Facility Requirements
- Cost Model
- Analysis of Options
- Summary
- Supporting Data

Motivation

- **In the commercial world, large scale computing is no longer optional (www.compete.org)**

"Nearly 100% of the respondents indicated that HPC tools are indispensable, stating that they would not exist as a viable business without them or that they simply could not compete effectively. A majority (70%) of the respondents indicated that HPC is so important that their organizations could not function without it."

- **This is even more true for the research community**

- **MIT researchers have had the ability (i.e. research $) to do it on their own for a long time**
  - **Much longer than most of our peers**
  - **Perhaps "too" successful: no longer have a place to put the computing resources**

- **This is an industry wide phenomena**
  - **Facility needs are major concern across the community**

Study Goals

- Focus on long term needs (2010-2012) timeframe
  - Get out ahead of current urgent campus situation
  - This will allow near term “stop gap” campus measures to be formulated in the context of an overall plan
  - Timeframe is also consistent with when current Lincoln F1 facility will become oversubscribed
- Goal 1: Develop a clear set of 2010+ requirements
- Goal 2: Develop a cost model for evaluating options
- Goal 3: Apply cost model to current location options
- Goal 4: Provide a reduced set of location options that can be analyzed more deeply

How Did I Get Here

- Member of Embedded Digital Systems Group (102)
  - Provides high performance computing solutions for the DoD
  - Conducts ~10 computer sizing studies a year
    - Increasing looking at larger scale system
- Lead architect LLGrid
  - 1500 Processor capability for Lincoln Staff (major F1 tenant)
- Performed F1 Cost Benefit Analysis
  - Have been thinking about the follow on for the last 2 years
- Lead DARPA High Productivity Computing Systems (HPCS) Productivity Team
  - Developing cost models of high performance computing for the biggest gov't users: DOE, DoD, NSA, NRO, NASA, …
  - Lots of info from centers regarding cost/benefit analysis

- Am I an expert? No
- Are there any experts? Not really, local conditions tend to be the most important in making these decisions
- Most facilities do this in an ad hoc manner once a decade
- We do have the benefit of their collective experience

Summary of Location Options

- Rental space within commuting distance of campus
  - Maximum convenience to campus
  - Minimum up front costs and risk
  - Maximum long term costs and risk
- MIT owned space within commuting distance of campus (Bates Lab, Lincoln Lab)
  - Still convenient
  - Maximum up front costs
  - Medium long term costs and risk
- Space beyond commuting distance of campus
  - Suggested by BFSG
  - Trade distance for costs
  - Medium up front costs with some risk
  - Minimum long term costs

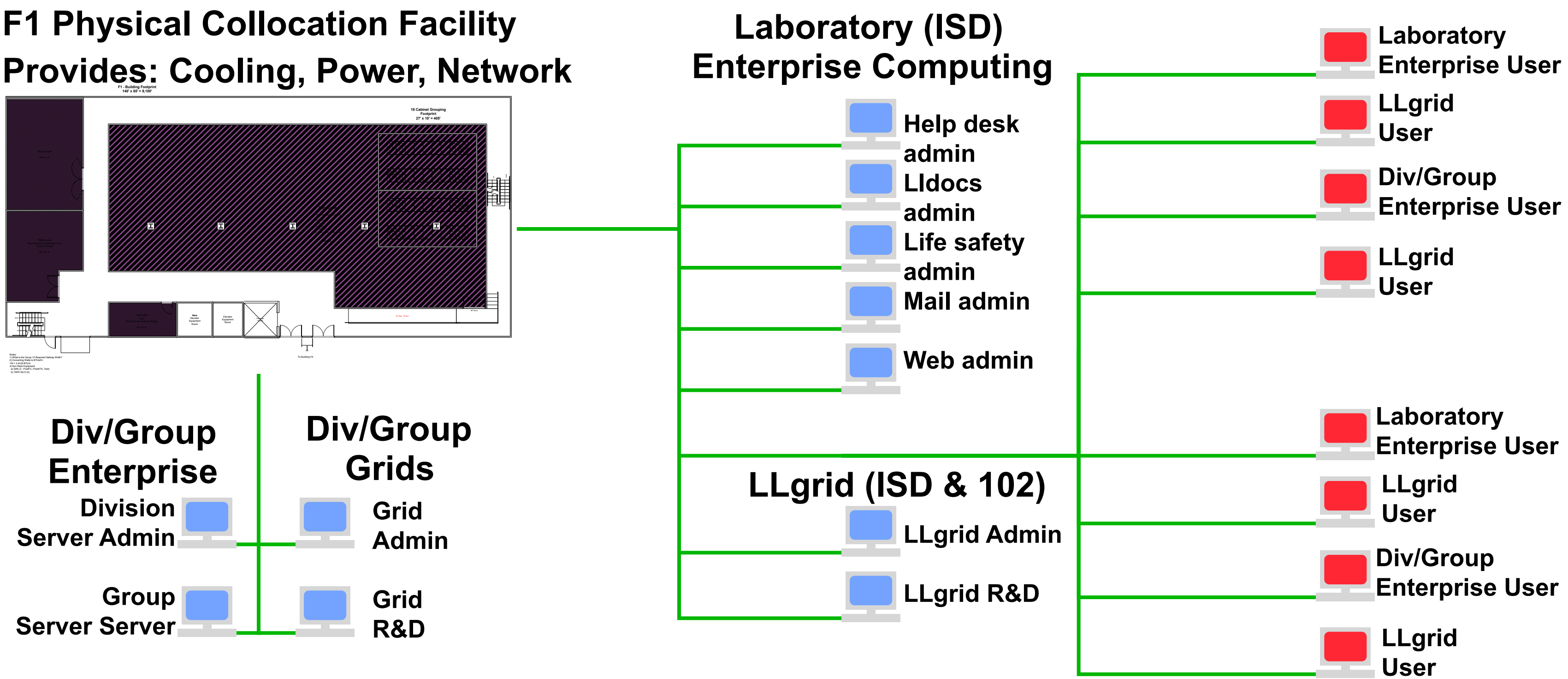
Outline

- Introduction
- **Facility Requirements**
  - ***What is it***
  - *Campus 2010+ needs*
  - *Lincoln 2010+ needs*
  - *Summary of needs*
- Cost Model
- Analysis of Options
- Summary
- Supporting Data

MIT Data Center Summary

- What: *Functionally Distributed* - Physically Collocated Computing
  - Admin, Grid, Divisions and Groups purchase, control and maintain their own systems (i.e. not centralized computing)
- Why: Fundamentally Enables MIT Mission
  - Enterprise computing made more reliable and easier to admin
  - Can address researchers who require computing for
    - Analyzing larger data sets
    - Modeling larger systems
- Will save significant $ over current ad hoc growth
  - Current and planned hardware sharing sub-par space and will require power and cooling upgrades
  - MIT Data Center provides required physical infrastructure with room for growth at less cost
  - Allows Gov't/Vendor funded computing equipment

Example: Lincoln F1 Data Center



Key

- Admin
- User
- LLAN

- *Functionally Distributed* - Physically Collocated Computing
  - ISD, Grid, Divisions and Groups purchase, control and maintain their own systems using LLAN
  - Users use systems from their desktops over LLAN
- Data Center provides physical infrastructure that would otherwise be provided in many lab spaces scattered across the Laboratory

Outline

- Introduction
- **Facility Requirements**
  - *What is it*
  - ***Campus 2010+ needs***
  - *Lincoln 2010+ needs*
  - *Summary of needs*
- Cost Model
- Analysis of Options
- Summary
- Supporting Data

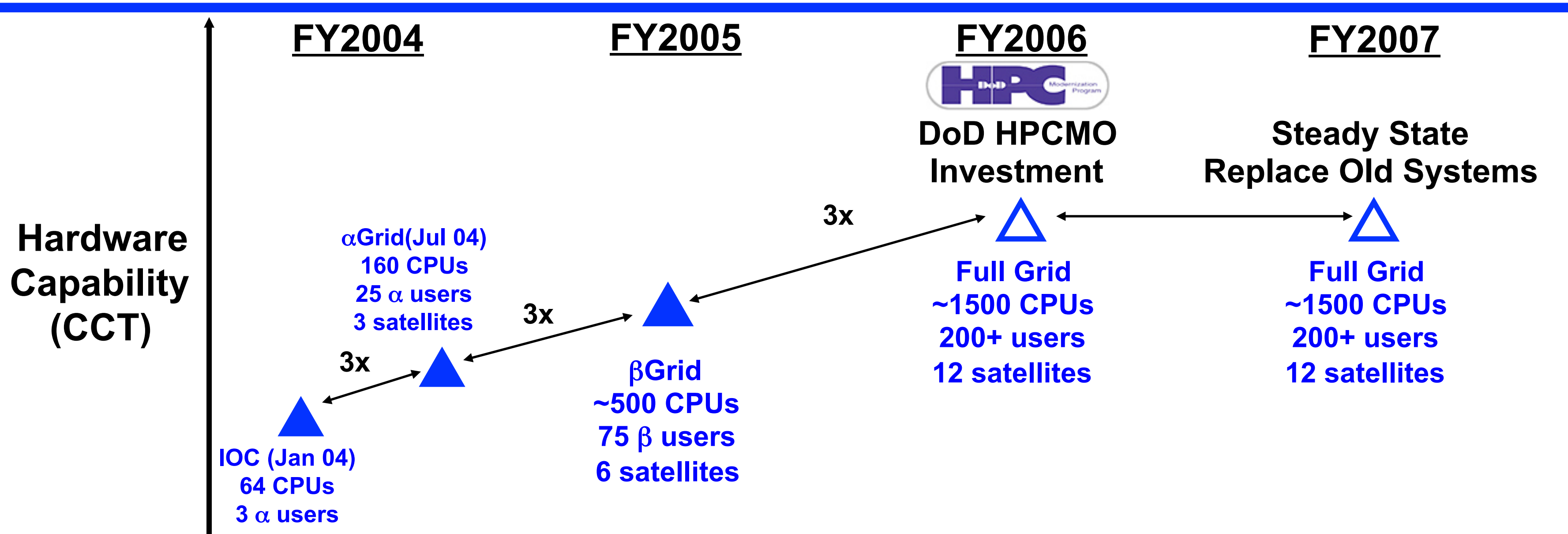
Campus Needs

- Currently ~100 racks of equipment in sub-par space distributed across campus
- Significant increase at rate at which equipment is arriving (due to relative decreases in hardware costs)
- ~500 individual labs across campus
  - By 2010 is not unreasonable to assume that 10% would acquire their own 100 processor compute cluster => ~5000 processor aggregate capability
  - By 2020 it is not unreasonable that both use and capability needs would increase by ~3x

Outline

- Introduction
- **Facility Requirements**
  - *What is it*
  - *Campus 2010+ needs*
  - ***Lincoln 2010+ needs***
  - *Summary of needs*
- Cost Model
- Analysis of Options
- Summary
- Supporting Data

Strategic Roadmap



	FY2004	FY2005	FY2006	FY2007
Software (102)	pMatlab+gridMatlab (accessibility)	pMatlab+petascale (applicability+maintainability)	pMatlab+DCT[1] 2.0 (scalability+reliability) XVM, pMapper to users	pMatlab+DCT[1] 3.0 pMapper for HPEC Classified Grid
Support (CCT+102)	Consulting Tutorial	Help Desk Class	Vendor Support Help Desk/ Classes Lab Rollout	Vendor & MIT Support Classified Grid
Admin (CCT)	Acquisition, OS, Network	Standard Cluster Image, F-Building, File System	Full Admin Support	Vendor Admin Support MIT-wide Solution

- **Key Challenges**
  - Management of 1000 processor system
  - Scaling pMatlab, XVM & pMapper to 1000 processors

[1] DCT = MathWorks Distributed Computing Toolbox

Lincoln Strategic Roadmap

	FY2006	FY2007	FY2008
Hardware Capability (CCT)	DoD HPCMO Investment Full Grid ~1500 CPUs 200+ users 12 satellites	Steady State Replace Old Systems Full Grid ~1500 CPUs 200+ users 12 satellites	LLGrid w/campus ~5000 CPUs (100 CPUs/grp) 400+ users
Software (102)	pMatlab+DCT[1] 2.0 (scalability+reliability) XVM, pMapper to users	pMatlab+DML[1] 3.0 pMapper for HPEC Classified Grid	HPCS Next Gen. Prog. Env. Adapted Multichan. Wideband Sig. Proc.
Support (CCT+102)	Vendor Support Help Desk/ Classes Lab Rollout	Vendor & MIT Support	Vendor & MIT Support
Admin (CCT)	Full Admin Support	Vendor Admin Support	Vendor Admin Support MIT-wide Solution

- **Key Challenges**
  - Management of 5000 processor system
  - Power, cooling, and space

[1] DCT = MathWorks Distributed Computing Toolbo

Lincoln Needs

- **Currently ~100 racks of equipment heading toward F1 data center**

- **~50 groups with ~50 individuals**
  - **By 2010 is not unreasonable to assume that 50% group would require their own 200 processor compute cluster => ~5000 processor aggregate capability**
  - **By 2020 it is not unreasonable that both use and capability needs would increase by ~3x**

Outline

- Introduction
- **Facility Requirements**
  - *What is it*
  - *Campus 2010+ needs*
  - *Lincoln 2010+ needs*
  - ***Summary of needs***
- Cost Model
- Analysis of Options
- Summary
- Supporting Data

Combined Physical Needs

Assumptions

- Current processor $ per form factor remain constant
- Large increases in per processor power requirements
- Future capability improvement primarily due to increased parallelism (scale out vs scale up)

2010 Requirements

- ~10,000 processor capability split ~50/50 between campus and Lincoln
- Assuming ~1 CPU/1U => 250 racks
- 30 sq. ft. rack => 7500 ft^2 of machine room floor (10,000 ft^2 of building)
- Power ~500 W/processor => ~5 MW to floor (~10 MW to building)

2020 Growth

- 3x 2010 Requirements (750 racks, 30 MW)
- Now only need an option for more space (30,000 ft^2)

- **A large, but not atypical facility**

Personnel Requirements

- Most admin personnel remain paid for out of groups that buy hardware
  - More groups may choose to leverage central admin capabilities
- Admins can remain where there are
- Number of personnel who need to be at site ("server huggers") is small (~10)
  - Monitoring physical state of system
  - Facility security
  - Assisting vendors with physical setup/repair
- Reducing physical access has been shown to significantly increase reliability
  - Fewer accidental touches of hardware

Outline

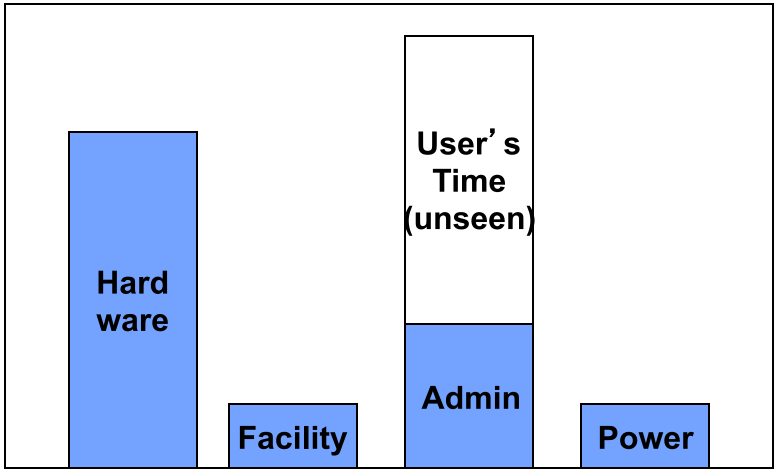
- Introduction
- Facility Requirements
- **Cost Model**
  - *Overview*
  - *Up Front Costs*
  - *Recurring Costs*
- Analysis of Options
- Summary
- Supporting Data

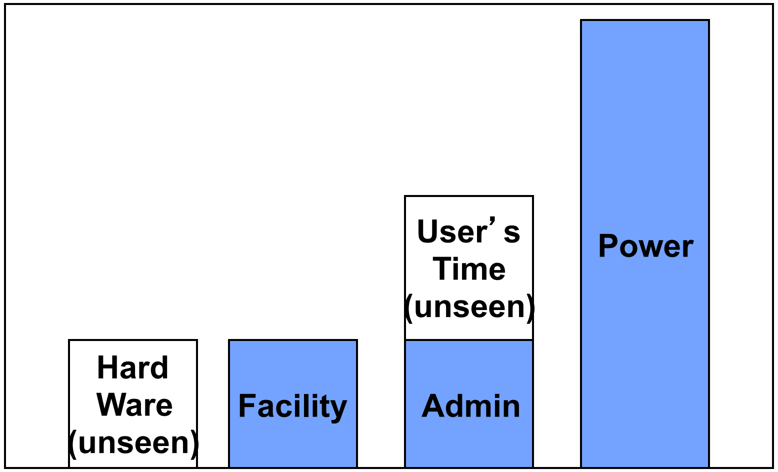
Cost Overview

- **Up Front Costs**
  - Facility setup (lasts 20 years)
    - Building Construction/Refurbishment, Network connection, Data Center Installation (Air Condition, Generators/Backup)
  - Compute Hardware (lasts 5 years)
    - Computers
    - Compute infrastructure Racks, Cabling, KVM switches, …
- **Recurring costs**
  - Personnel
    - Sys Admins
    - Building Personnel
    - Users time
  - Electrical Power

Relative Costs Over 10 Years

- **Traditional View**
  - Hardware cost is dominant (cheap clusters change this)
  - Facility is cheap (increasing size needs change this)
  - User's Time is unseen (user's time is critical)
  - Bigger staff is better (fight server hugging)
  - Power is unimportant (increasing size needs change this)
- **MIT Reality**
  - Clusters are cheap and cost is born by research groups
  - Facility is expensive to setup, but cheap to maintain
  - Cluster admin is rapidly improving
  - Electrical costs are big and increasing





Generic Facility Up Front Costs

- Facility setup (every 20 years)
  - Building Construction $\$500/ft^2 \times 10{,}000\ ft^2 = \$5M$ (50% labor)
    - Based on Lincoln Annex 1 construction costs
  - Refurbishment (site specific)
  - Network connection $100K/mile x 10 miles = $1M
  - Data Center installation $\$1000/ft^2 \times 7{,}500\ ft^2 = \$7.5M$ (50% labor)
    - Based on Lincoln F1 construction costs
- Compute Hardware (every 5 years)
  - Computers $1K/processor x 10,000 processors = $10M
    - Paid for by groups
  - Rack infrastructure $10K/rack x 250 rack = $2.5M
    - Based on F1 construction costs

Generic Facility Recurring Costs

- **Personnel**
  - Sys Admins $200K/yr x 10 = $2M/yr
    - Paid for by departments or labs
  - Building Personnel $200K/yr x 10 = $2M/yr
  - User's Time 1% x $200K/yr x 1000 users = $2M/yr
    - Paid for by departments or labs
    - Domain specific solutions keep this cost down
- **Power**
  - Electrical Power 10 MW/yr x $1M/(MW/yr) = $10M/yr
    - Cogeneration not suitable/beneficial?

Generic Facility 10 Year Costs

- Facility $13.5M
- Rack Infrastructure $5M
- Building Personnel $20M
- Electrical Power $100M

Outline

- Introduction
- Facility Requirements
- Cost Model
- **Analysis of Options**
  - *Rental*
  - *Bates*
  - *Lincoln*
  - *Other*
- Summary
- Supporting Data

Local Rental Space

- **Change from generic**
  - Facility setup costs covered by rent
  - $?/ft^2 x 10,000/ft^2 = $?
- **All other costs the same**
- **Net = +/- $?**
- **Advantages: speed, lower up front cost**
- **Risks: increased rent, may not be expandable, may not meet security requirements**

Bates Laboratory

- Change from generic: none?
  - Does a building already exist
- Net = $0
- Advantages
  - Physically convenient
- Disadvantages
  - Maximum up front costs
  - Medium long term costs and risk

- Change from generic
  - No cabling costs
- Net = -$1M
- Advantages
  - Physically convenient
  - Clearly meets security requirements
- Disadvantages
  - Maximum up front costs
  - Medium long term costs and risk

Space Beyond Commuting Distance

- Idea suggested by BFSG
  - Spirit of suggestion: can we trade distance for cost
  - Two locations in the state with the following benefits
    (1) Labor costs are 1/2; (2) Power costs are 1/4
- Cost impact
  - Building Construction = -$1M
  - Data Center installation = -$1.5M
  - Building Personnel = -$1M/yr
  - Power = -$7.5M/yr
  - Net 10 year impact = -$87.5M
- Net 10 Year Impact = -$87.5M

Current MIT Electrical Power

- Cogeneration facility (http://cogen.mit.edu/powermit/)
  - Furnace that happens to generate electricity ~20 MW
  - Excellent solution during winter time
  - As a pure electrical generator it is expensive ~$2M/(MW/yr)
  - Power only available to facilities on campus?
  - Data Center won't need heat

Current MIT Power

This page automatically refeshes in 00:11 seconds

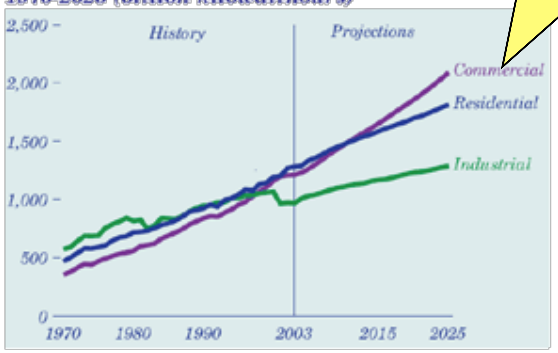
Instantaneous Power	Present (MegaWatts)	Today (MegaWattHours)	Present Cost* ($/s)
Current MIT Demand	22.006	185.524	1.13
Produced by MIT 20 MW Turbine	21.900	186.007	1.13
Produced by NSTAR	0.306	1.370	0.01

*MIT price based on fuel flow (natural gas and distillate oil), and price of natural gas ($4.60 / thousand cubic feet) and distillate ($1.59 / gallon) Energy Information Administration, 2002. NSTAR price based on NSTAR load and NSTAR default service price for large commercial service ($58.77 / MWH) NSTAR Default Sevice, 2003.

"Projected efficiency gains for electric equipment in the commercial sector are offset by the continuing penetration of new telecommunications technologies and medical imaging equipment, increased use of office equipment, and more rapid additions of floorspace."

- Utility Power (http://www.eia.doe.gov)
  - Campus pays $600K/(MW/yr)
  - Lincoln pays $800K/(MW/yr)
  - Both are expected to increase

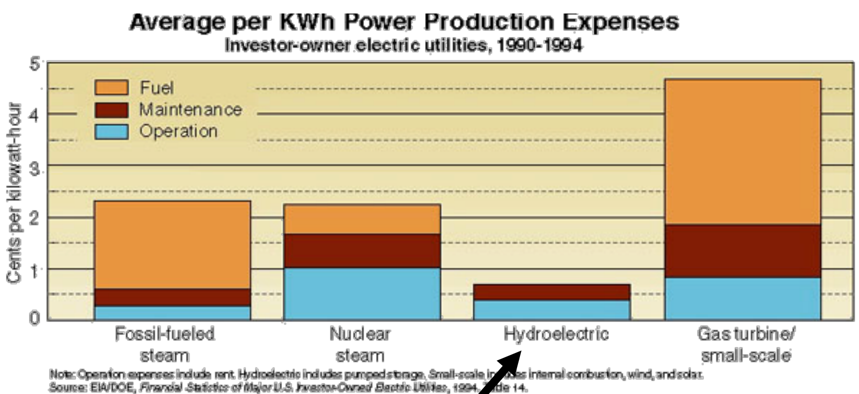
Figure 66. Annual electricity sales by sector, 1970-2025 (billion kilowatthours)



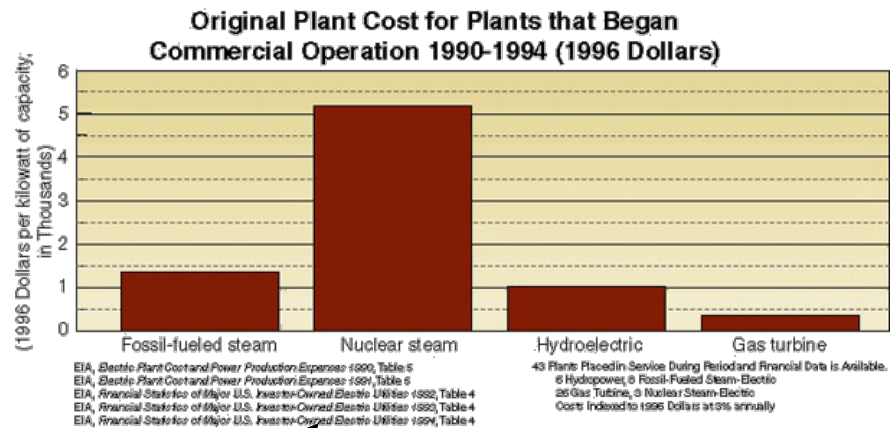
How to find cheap power

- Source http://hydropower.inel.gov/hydrofacts/plant_costs.shtml
- 1994 data has been validated against more recent cost estimates

Cost to operate



Cost to build



- Hydroelectric is cheapest to operate
- Look for a facility that is already "paid for"
- Look for a smaller facility that would view us as a big customer
- Look for a community that is interested in ancillary benefits
  - Investment, jobs, …

Distribution of Hydro Power in US

- Source http://hydropower.inel.gov/resourceassessment/pdfs/03-11111.pdf
- Source http://hydropower.inel.gov/resourceassessment/states.shtml

Hydro Resources by Region

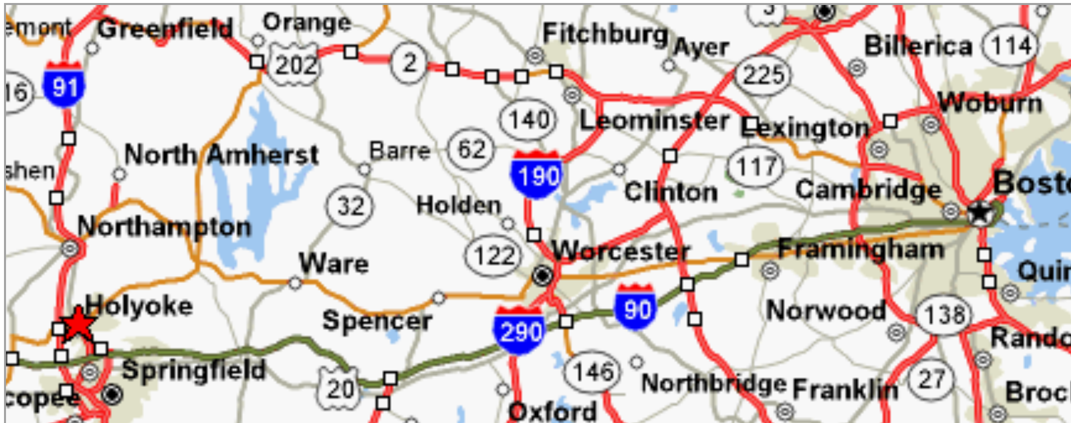
Region (HUC)	Name	Average Annual Mean Power (Developed Potential) (MW)	Average Annual Generation (MWh)	Developed Capacity (MW)	Number of Plants
1	North Atlantic	873	7,648,300	1,881	397
2	Mid-Atlantic	840	7,359,758	2,060	206
3	South Atlantic-Gulf	1,849	16,195,298	6,743	165
4	Great Lakes	2,852	24,986,998	4,092	288
5	Ohio	820	7,182,482	1,772	48
6	Tennessee	1,859	16,282,814	3,855	55
7	Upper Mississippi	404	3,540,641	734	119
8	Lower Mississippi	136	1,192,680	398	6
9	Souris Red-Rainy	13	110,058	22	8
10	Missouri	1,797	15,743,664	3,722	80
11	Arkansas-White-Red	696	6,100,625	2,097	33
12	Texas Gulf	127	1,115,557	428	23
13	Rio Grande	50	441,821	157	7
14	Upper Colorado	724	6,339,303	1,882	41
15	Lower Colorado	789	6,911,489	2,556	23
16	Great Basin	97	853,413	228	81
17	Pacific Northwest	16,645	145,811,168	32,365	339
18	California	4,668	40,892,958	9,450	413
19	Alaska	171	1,500,596	392	40
20	Hawaii	20	173,300	38	16
	Totals	**35,432**	**310,382,923**	**74,872**	**2,388**

- New England has many small hydro facilities
- All are "paid for"
- Holyoke, MA and Montague, MA are the largest (~50 MW each)

Largest Sites in MA

Station	River	(Kilowats)	Units	Location
Beebe-Holbrook	Connecticut	400	2	Holyoke, MA
Boatlock	Connecticut	2,900	3	Holyoke, MA
Chemical	Connecticut	1,500	2	Holyoke, MA
Hadley Falls	Connecticut	31,500	2	Holyoke, MA
Riverside	Connecticut	6,900	4	Holyoke, MA
Skinner	Connecticut	260	1	Holyoke, MA
Turners Falls	Connecticut	6,000	5	Montague, MA
Cabot	Connecticut	53,000	6	Montague, MA
Cobble Mountain	Little River	33,000	3	Granville, MA

Holyoke, MA (Pop. 40,000)

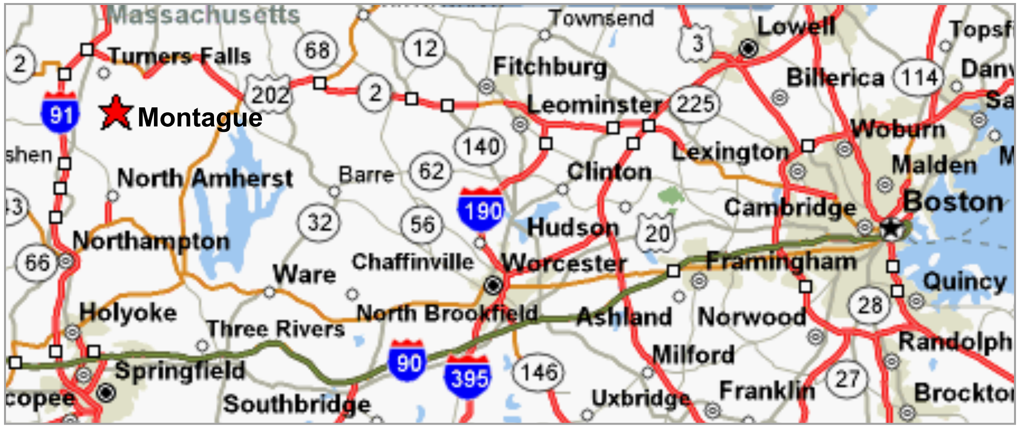


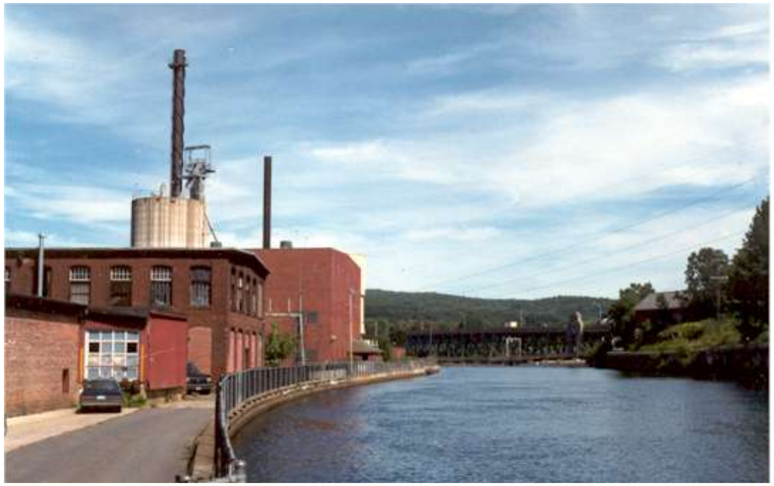
Site Info
- Motto "The Power to Grow"
- Holyoke Gas & Electric (City)
- Cheap power $250K/(MW/yr)
- Dependable power (100+ years)
- Lots of space for facility

Labor Costs
- Median household income $30K
- Median house cost $105K

Local Institutions
- Holyoke Community College
- Close to U.Mass. & Five Colleges

Montague, MA (Pop. 8,500)





Site Info
- For sale by Northeast Utilities
- Cheap power $250K/(MW/yr)?
- Dependable power (100+ years)
- Lots of space for facility

Labor Costs
- Median household income $34K
- Median house cost $112K

Local Institutions
- Greenfield Community College
- Close to U.Mass. & Five Colleges

Miscellaneous

- Communities are looking for economic opportunities
  - Potential local, state, federal support for project?
- Potential to work with local community colleges
  - Training of on site personnel?
- Potential to work/collaborate with U. Mass. & Five Colleges
- Network distance 10 miles?

- Driving Distance
  - Both near major roads (Mass Pike and Route 2)
  - Lincoln to Stata takes 50 minutes
  - Lincoln to Holyoke takes 80 minutes
  - Lincoln to Montague takes 90 minutes
  - Cambridge to Holyoke takes 80 minutes
  - Cambridge to Montague takes 100 minutes

This is not new

- SUNY Buffalo Computing Center (1500 processors)
- Exploits cheap local power

- Creating a low cost, high performance facility is not difficult
- The real challenge is having the talented scientists and engineers who can turn it into great research
  - MIT has this much rarer commodity in abundance

Summary

- **Compute resources are core to MIT mission**
  - Have outgrown current facility
- **In 2010 timeframe will require**
  - 10,000 processors, 750 racks, 10,000 ft^2, 10MW
- **Generic facility costs**
  - $18.5M to build, $2M/yr personnel, $10M/yr power
- **Analyzed several facility sites**
  - Diverse range of cost options