IBM High Performance Computing in a Cloud Computing Model

Cluj-Napoca, 25.10.2012

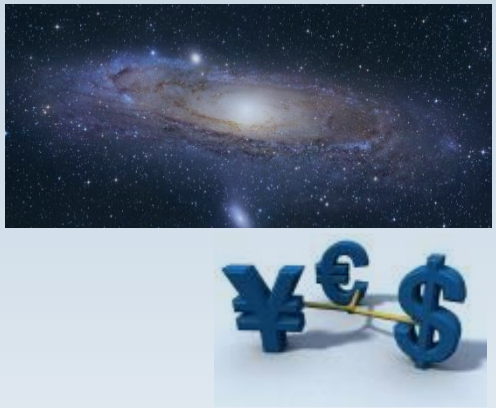
Florin MANAILA
Deep Computing Architect - IBM South East Europe
florin.manaila@ro.ibm.com
www.ibm.com/technicalcomputing

IBM Technical Computing – For Science and Business

Delivers Faster, More Reliable, More Responsive Results

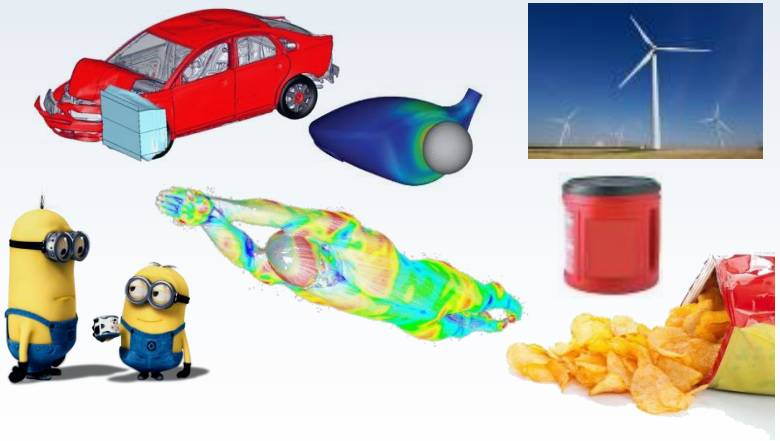
Technical Computing is how:

- The universe is analyzed
- Medical research is conducted
- Financial risk analysis is done
- Weather is forecasted
- Oil is discovered



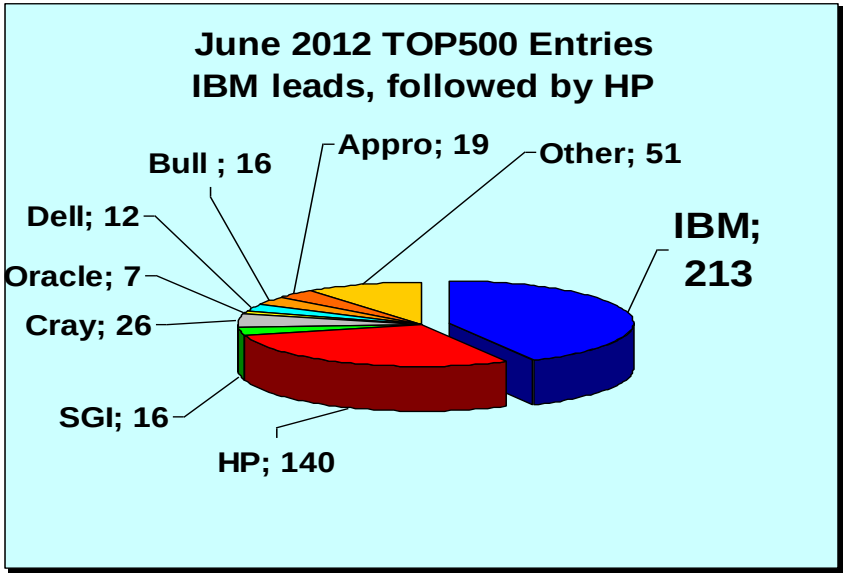
It's also how:

- Cars are designed and crash tests are simulated
  - Wind power is deployed to maximize output
  - Bathing suits are designed for minimal drag
- Golf clubs are developed to increase distance
  - Potato chips and coffee are packaged
  - Animated movies are made

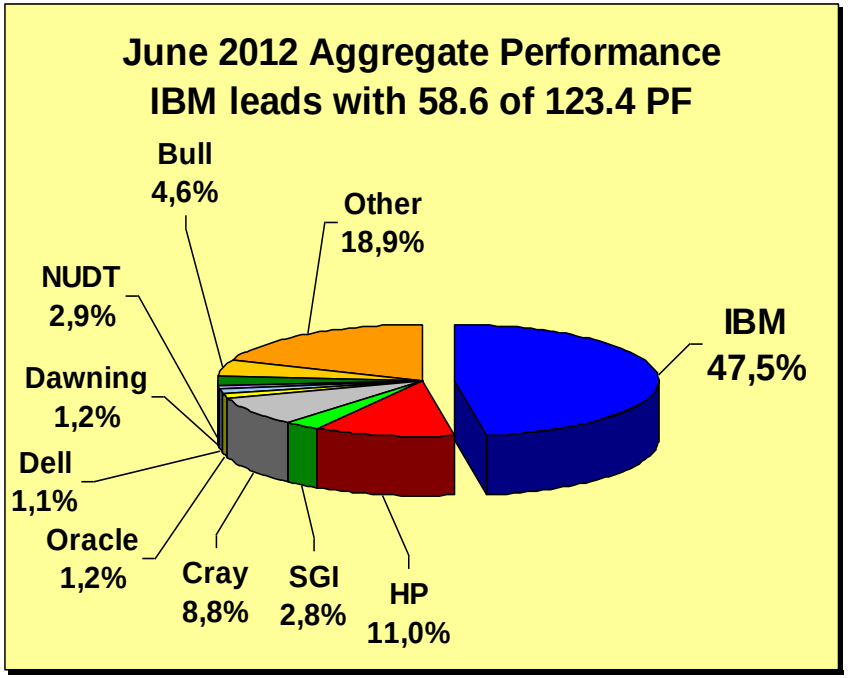


Technical Computing is how organizations achieve insight faster, reduce time and cost, and gain competitive advantage

IBM supercomputing leadership



Semiannual independent ranking of the top 500 supercomputers in the world



Source: www.top500.org

IBM supercomputing leadership ...

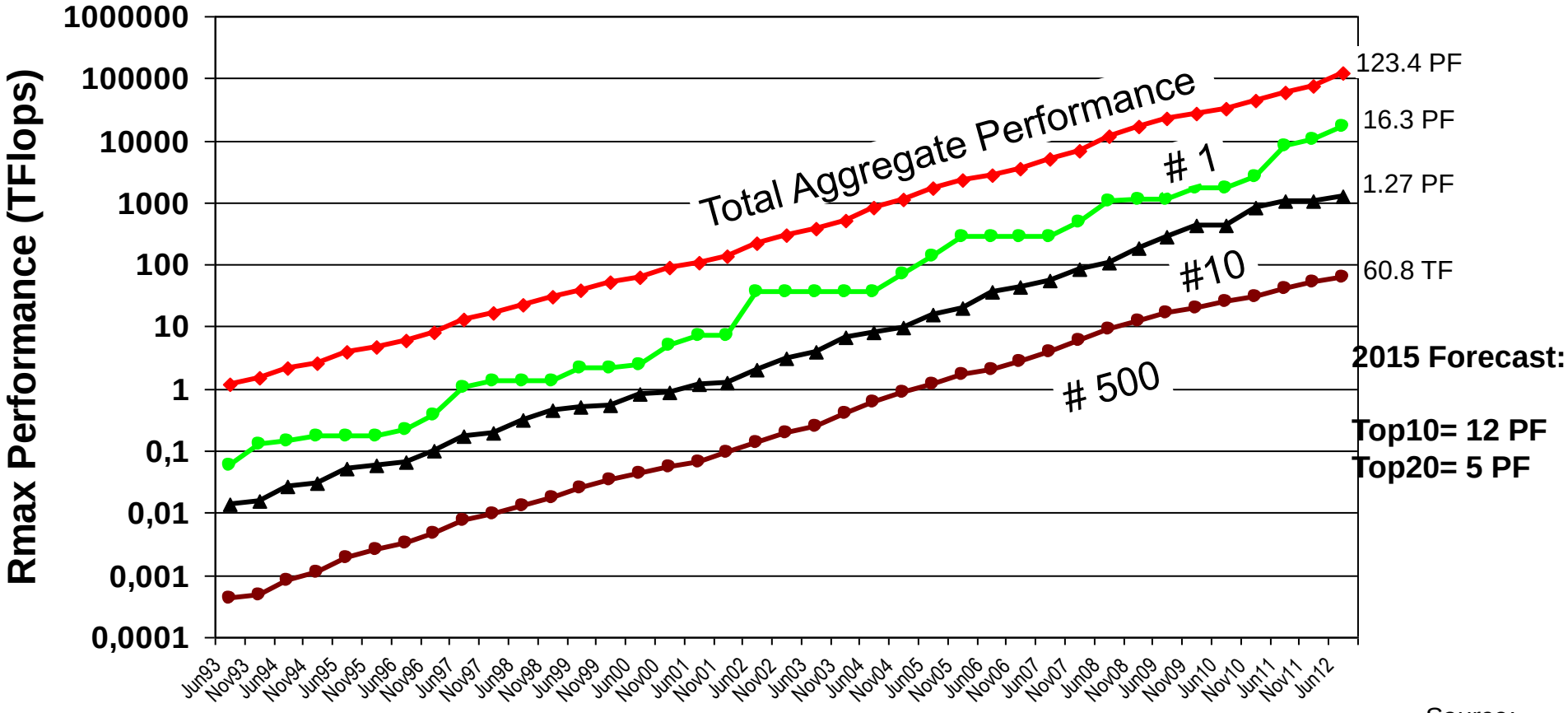
- ✓ New #1: LLNL Sequoia Blue Gene/Q 16.32 PFlops
  - ✓ 5 in TOP 10 (#3 ANL-Mira BG/Q, #4 LRZ-SuperMUC iDataPlex, #7 CINECA-Fermi BG/Q, #8 Juelich-JuQUEEN BG/Q)
- ✓ Most installed aggregate throughput with over 58.6 out of 123.4 Petaflops (47.5%) *(HP 13.5/11%, Fujitsu 12.2 /9.9%, Cray: 10.9/ 8.8%)*

  Lead for 26 Lists in a row
- ✓ Most systems in TOP500 with 213 *(HP: 140, Cray: 26, Appro: 19)*
- ✓ Fastest system in Europe (LRZ-SuperMUC iDataPlex)
- ✓ Fastest x86 system (LRZ-SuperMUC iDataPlex)
- ✓ **20 Most energy-efficient systems**

  All IBM Blue Gene/Q's

TOP500 Performance Trend

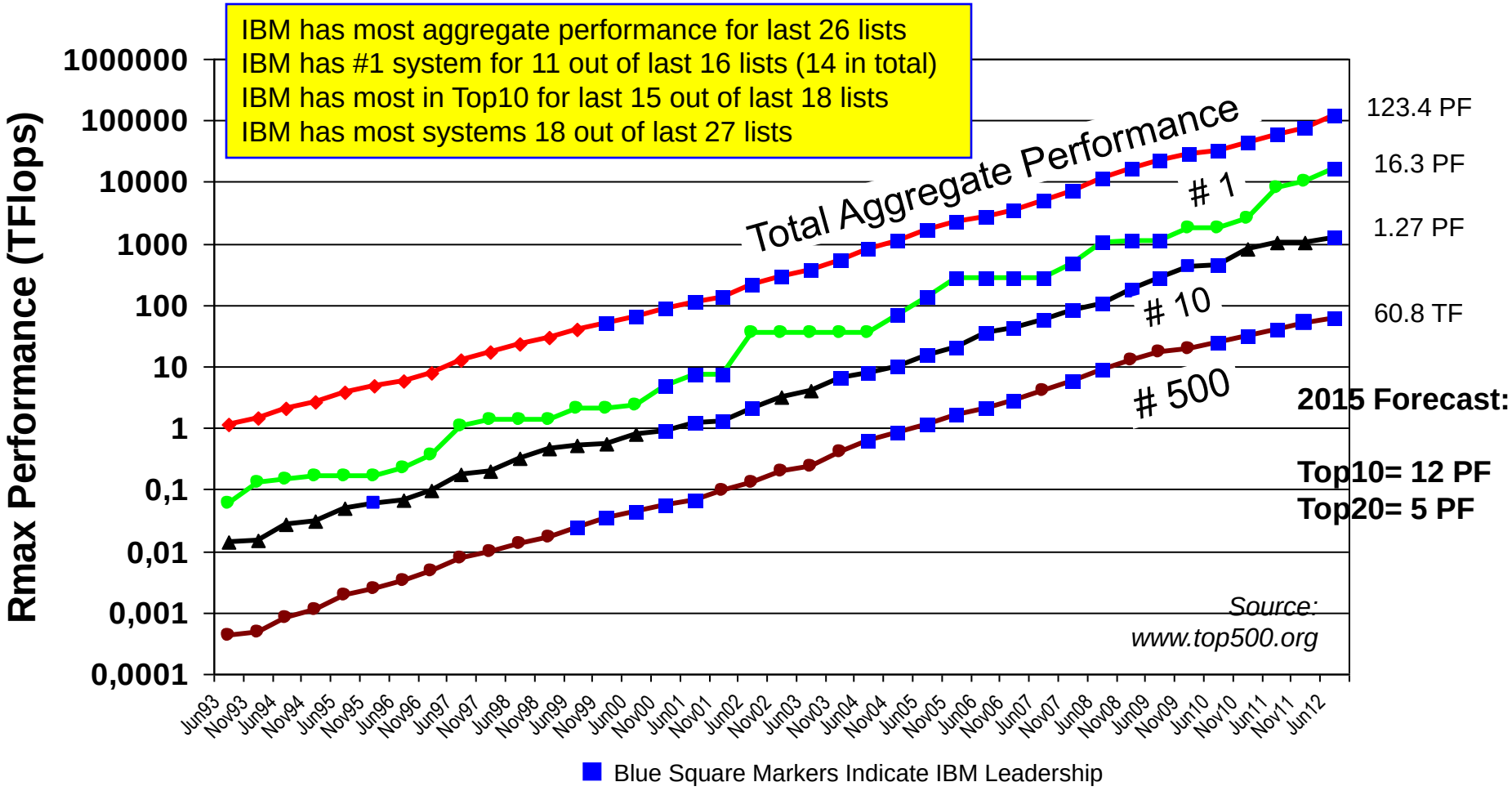
Even though there is some stepping of the performance of the #1 system, the relentless growth of the #500 clip level, #10 clip level and Total Aggregate performance all are fairly straight line trends when plotted on log scale.



Source: www.top500.org

TOP500 Performance Trend

Over the long haul IBM has demonstrated continued leadership in various TOP500 metrics, even as the performance continues it's relentless growth.



TOP20 Supercomputers

June 2012

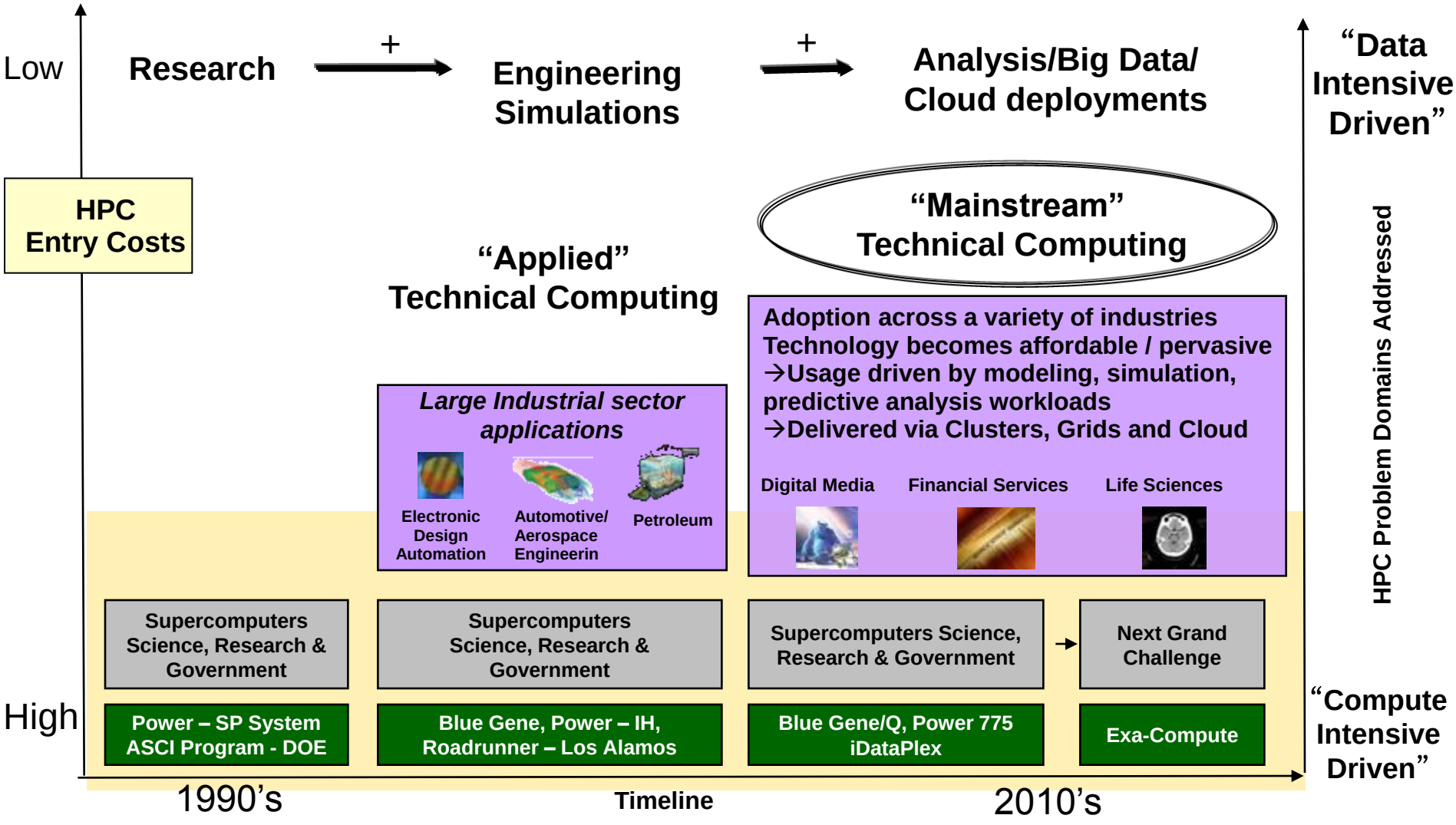
All Top20 Systems are over 1 PFlops.
IBM has most systems in TOP20 with eight.

#	Ven-dor	Rmax TFlops	Installation	
1	**IBM**	**16324**	**DOE/NNSA/LLNL-Sequoia** (96 racks Blue Gene/Q)	Update
2	Fujitsu	10510	**RIKEN K Computer – Japan** (SPARC64 VIIIfx 2.0GHz)	
3	**IBM**	**8162**	**DOE/NNSA/ANL - Mira** (48 racks Blue Gene/Q)	New
4	**IBM**	**2897**	**LRZ - SuperMUC** (iDataPlex – Sandy Bridge)	New
5	NUDT	2566	**NUDT – Tianhe 1A China** (Westmere/NVIDIA GPU)	
6	Cray	1941	**Oak Ridge NL - Jaguar** (XT5 6C 2.6 GHz Opteron)	Update
7	**IBM**	**1725**	**CINECA - Fermi** (10 racks Blue Gene/Q)	New
8	**IBM**	**1380**	**Juelich - JuQUEEN** (8 racks Blue Gene/Q)	New
9	Bull	1359	**CEA/GENCI - Curie** (Sandy Bridge)	New
10	Dawn-ing	1271	**NSCS-China Nebulae** (6C Xeon+NVIDIA GPU)	

#	Vendor	Rmax TFlops	Installation	
11	SGI	1243	**NASA Ames - Pleiades** (Altix mixed CPUs)	Update
12	Bull	1237	**IFERC Helios - Japan** (Sandy Bridge)	New
13	**IBM**	**1207**	**Daresbury – Blue Joule** (7 rack Blue Gene/Q)	New
14	HP	1192	**TiTech Tsubame 2.0–Japan** (Westmere/NVIDIA GPU)	
15	Cray	1110	**Sandia - Cielo** (XE6 8core 2.4GHz Opteron)	
16	Cray	1053	**NERSC – Hopper** (XE6 12core Opteron)	
17	Bull	1050	**CEA Tera-100 – France** (Nehalem-EX)	
18	Fujitsu	1043	**Univ Tokyo – Oakleaf FX** (SPARC64 IXfx 1.8 GHz)	New
19	**IBM**	**1042**	**DOE/NNSA/LANL - RR** (QS22/LS21)	
20	**IBM**	**1035**	**Univ Edinburgh – DiRAC** (6 rack Blue Gene/Q)	New

Source: www.top500.org

Pure science is expanding



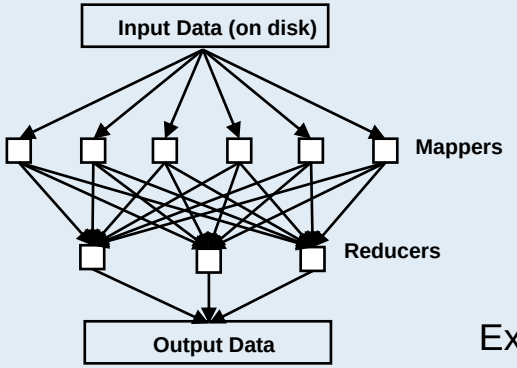
Styles of Massively Parallel Systems Underpin HPC and Analytics

Embarrassingly Parallel | Network Dependent

Random Comms

Data Intensive: Data at Rest

JAQL, Java



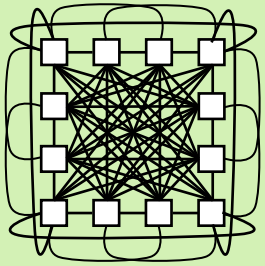
Data at Rest*:
- **High Volume**
- **Mixed Variety**
- **Low Velocity**

Extreme Scale-out

(*pre-partitioned)

Data and Compute Intensive (Large Address Space)

C/C++, Fortran, UPC, SHMEM, MPI, OpenMP



Data is Moving:
- **Long Running**
- **All Data View**
- **Small Messages**

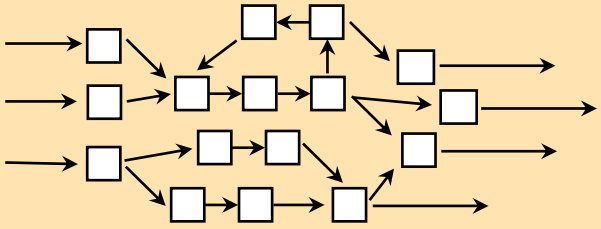
Discrete Math
Low Spatial Locality

Structured Comms

Data Intensive: Streaming Data

SPL, C, Java

Data in Motion:
- **High Velocity**
- **Mixed Variety**
- **High Volume***

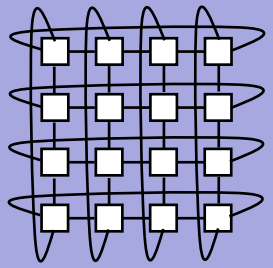


Reactive Analytics
Extreme Ingestion

(*over time)

Compute Intensive (Data Generators)

C/C++, Fortran, MPI, OpenMP



Data is Generated:
- **Long Running**
- **Small Input**
- **Massive Output**

= compute node

Generative Modeling
Extreme Physics

Industry Trends– what is happening in HPC?

- Exascale computing will be a reality in 2018/2019; large scale is being tackled now. Universities/research institutions typically will reach 10% of national lab installations. So in 2018, will large university installations have a multi petaflop computer? What will house it? What will be the power requirements? The Power Utilization Efficiency (PUE) of your datacenter is as important as the "green solution" you put in it.

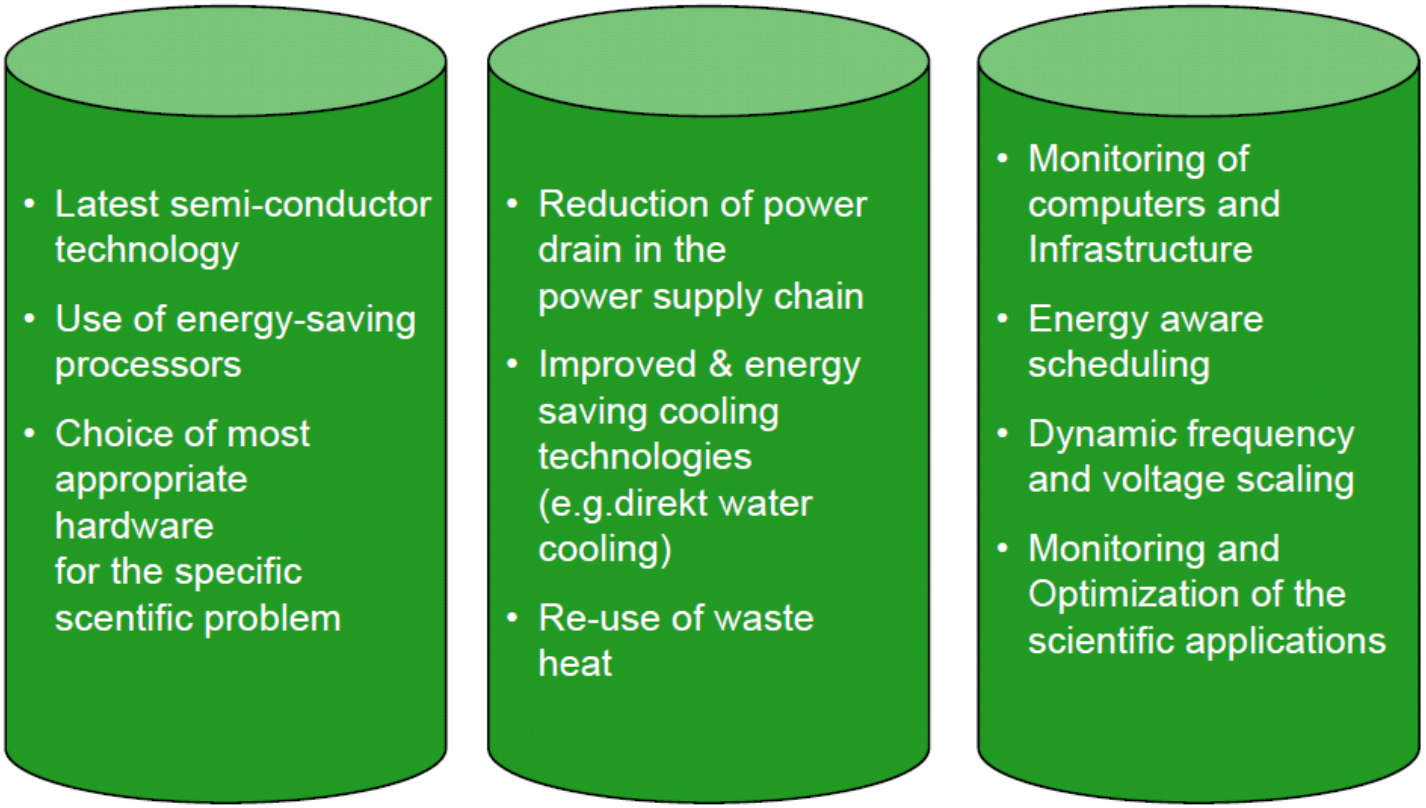
- The Data Deluge– big data, big data management is consuming researchers now. There are very large world-wide projects where data is measured in the 100s of petabytes. A petabyte is 10^{15} bytes of data. Big Insights, Infosphere Streams and Watson are important technologies from IBM that will drive the high end, and our file system (GPFS) will be an important differentiator.

- Justifying HPC on Campus (the role of the Cloud)– can be tackled by fulfilling the a myriad of campus needs for both high throughput computing and high performance (capability) computing using a shared environment. Best practices show that costs are reduced when one builds a central condominium facility where researchers can contribute their grant money and which serves the larger research community. HPC makes a research institution more competitive for grants and leads to more publications.

Key aspects in decisions:

- System Design
  - Project and design responsibility
- Data
- Economics
  - Environmental responsibility (GREEN): space, power, cooling
- Software and Software programming models
- Productivity
  - Management tools
- Reliability and availability

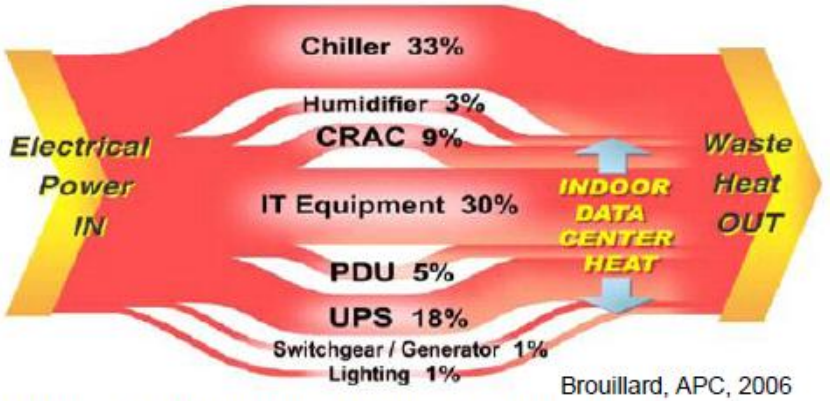
Pilars of Energy Efficient Computing



Green Datacenter Market Drivers and Trends

- Increased green consciousness, and rising cost of power
- IT demand outpaces technology improvements
  - Server energy use doubled 2000-2005; expected to increase15%/year
  - 15 % power growth per year is not sustainable
  - Koomey Study: Server use 1.2% of U.S. energy
- ICT industries consume 2% worldwide energy
  - Carbon dioxide emission like global aviation

Source IDC 2006, Document# 201722, "The impact of Power and Cooling on Datacenter Infrastructure, John Humphreys, Jed Scaramella"

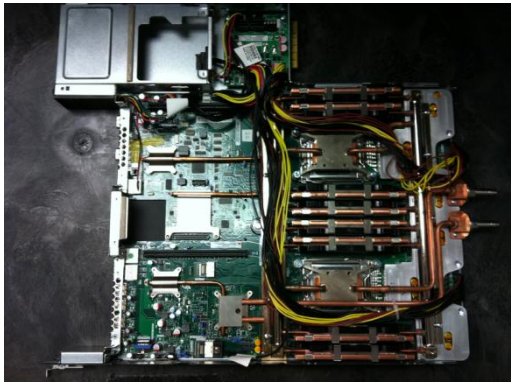


Brouillard, APC, 2006

Future datacenters dominated by energy cost; half energy spent on cooling

IBM Product Positioning

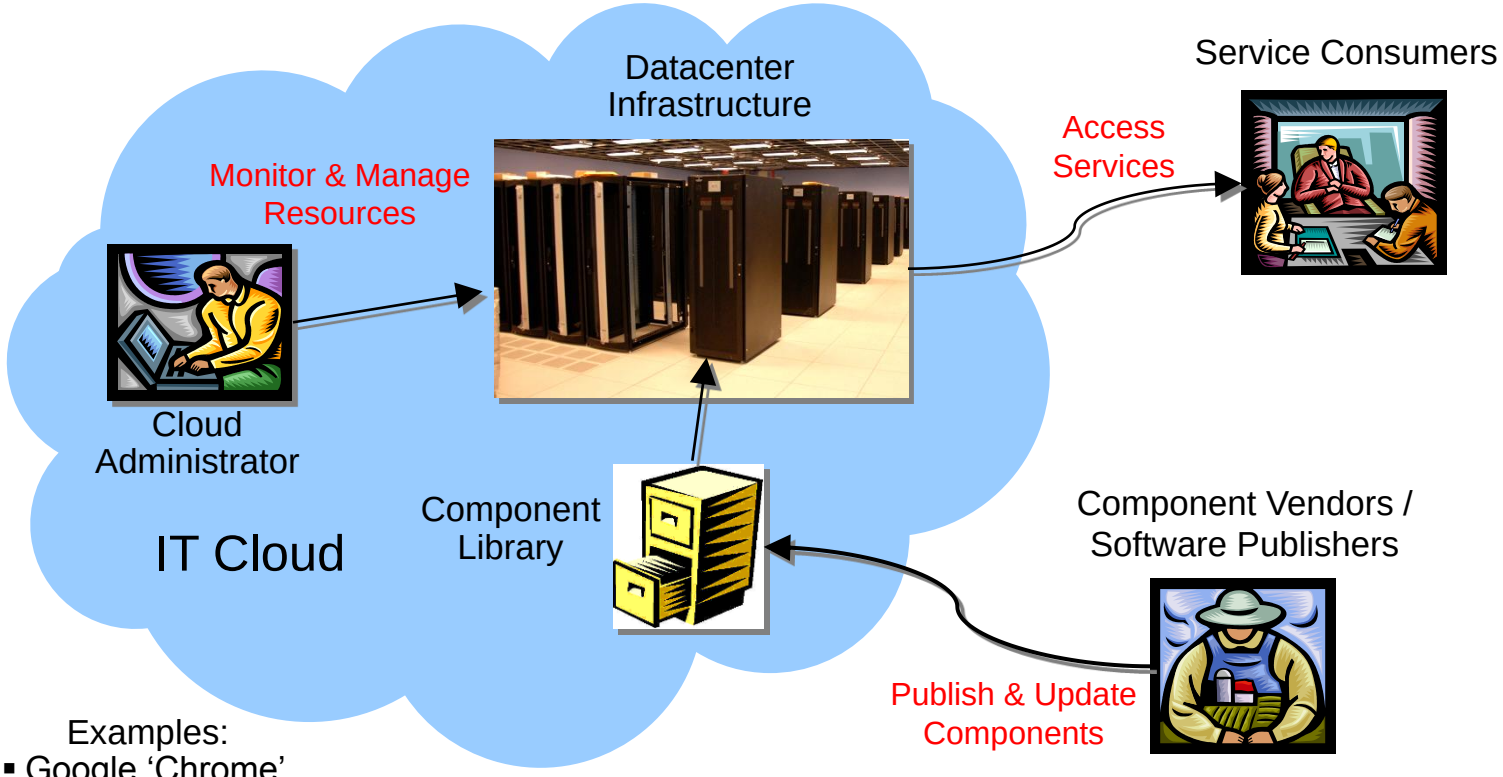
- **Blue Gene**
  - Ultra reliable
  - Ultra high scaling capability
  - Fast interconnect
  - Highly energy efficient
  - Very dense packaging
  - Strong PEAK $/Mflp price/performance
- **X86 Clusters**
  - Focused on “capacity”, scalability
  - High ISV coverage
  - Strong PEAK $/Mflp price/performance
  - GPU support
- **POWER Platforms**
  - Production ready, ultra reliable
  - Market leader – sustained application performan
  - Large memory SMP
  - Rich s/w stack (from PERCS)
  - Fast interconnect
  - Very dense packaging



HPC IN THE CLOUD

What Is Cloud Computing?

- Cloud computing is an emerging style of computing in which applications and data are provided as services to users over the Web.
  - ⇨ The services provided can be available globally, always on, low in cost, “on demand”, and massively scalable



Examples:
- Google ‘Chrome’
- Amazon Elastic Compute Cloud EC^2
- IBM Research Compute Cloud RC^2

Evolution of Cloud Computing

1980

Grid Computing

- Solving large problems with parallel computing
- Made mainstream by Globus Alliance

1990

Utility Computing

- Offering computing resources as a metered service
- Introduced in late 1990s

2000

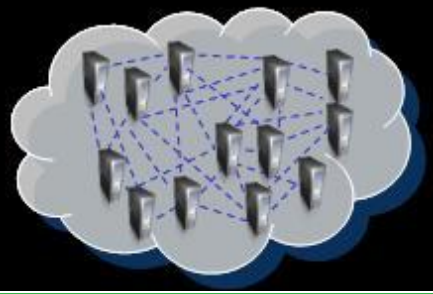
Software as a Service

- Network-based subscriptions to applications
- Gained momentum in 2001

2007

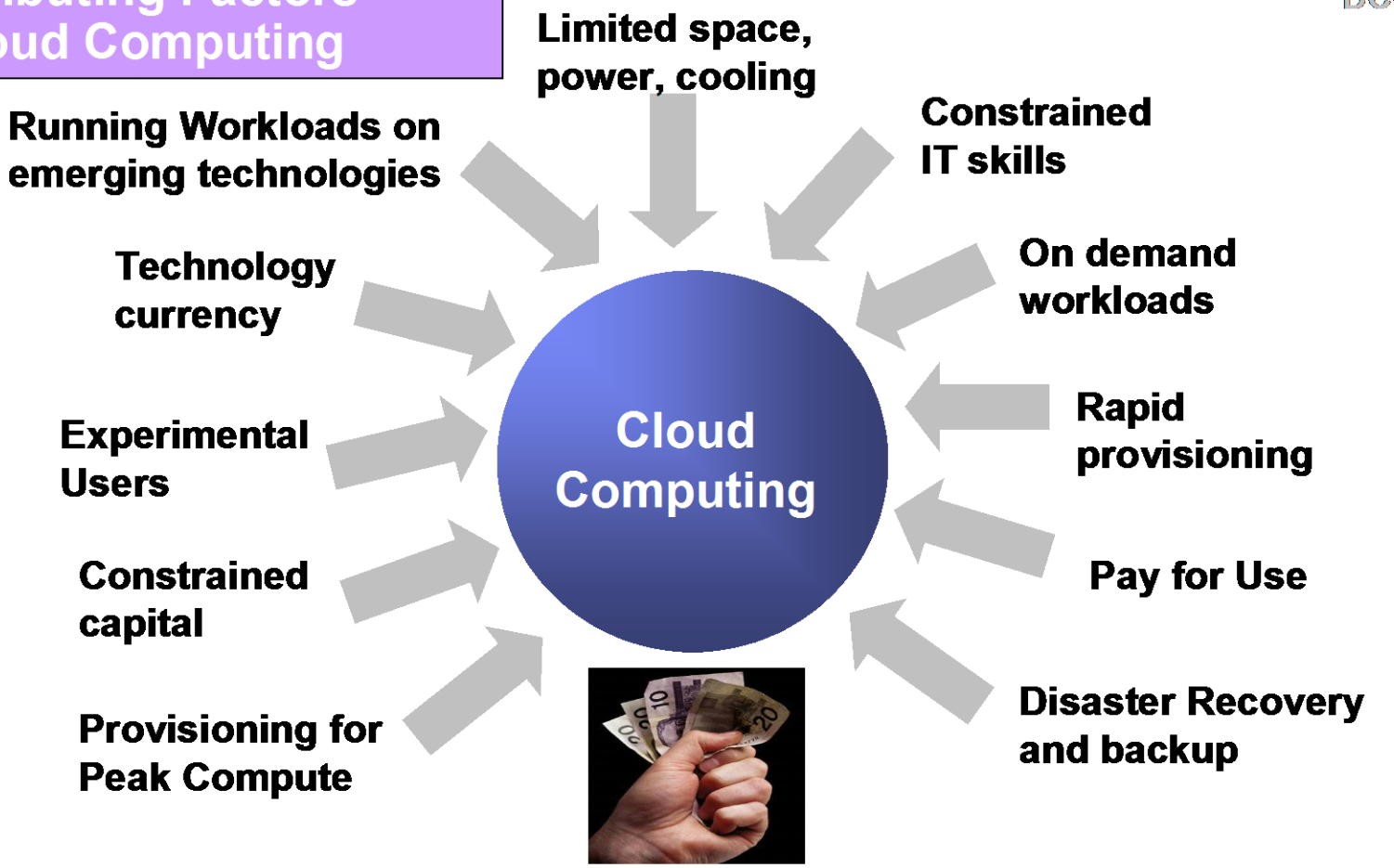
Cloud Computing

- Next generation Internet computing
- Next generation data centers



Contributing Factors to Cloud Computing

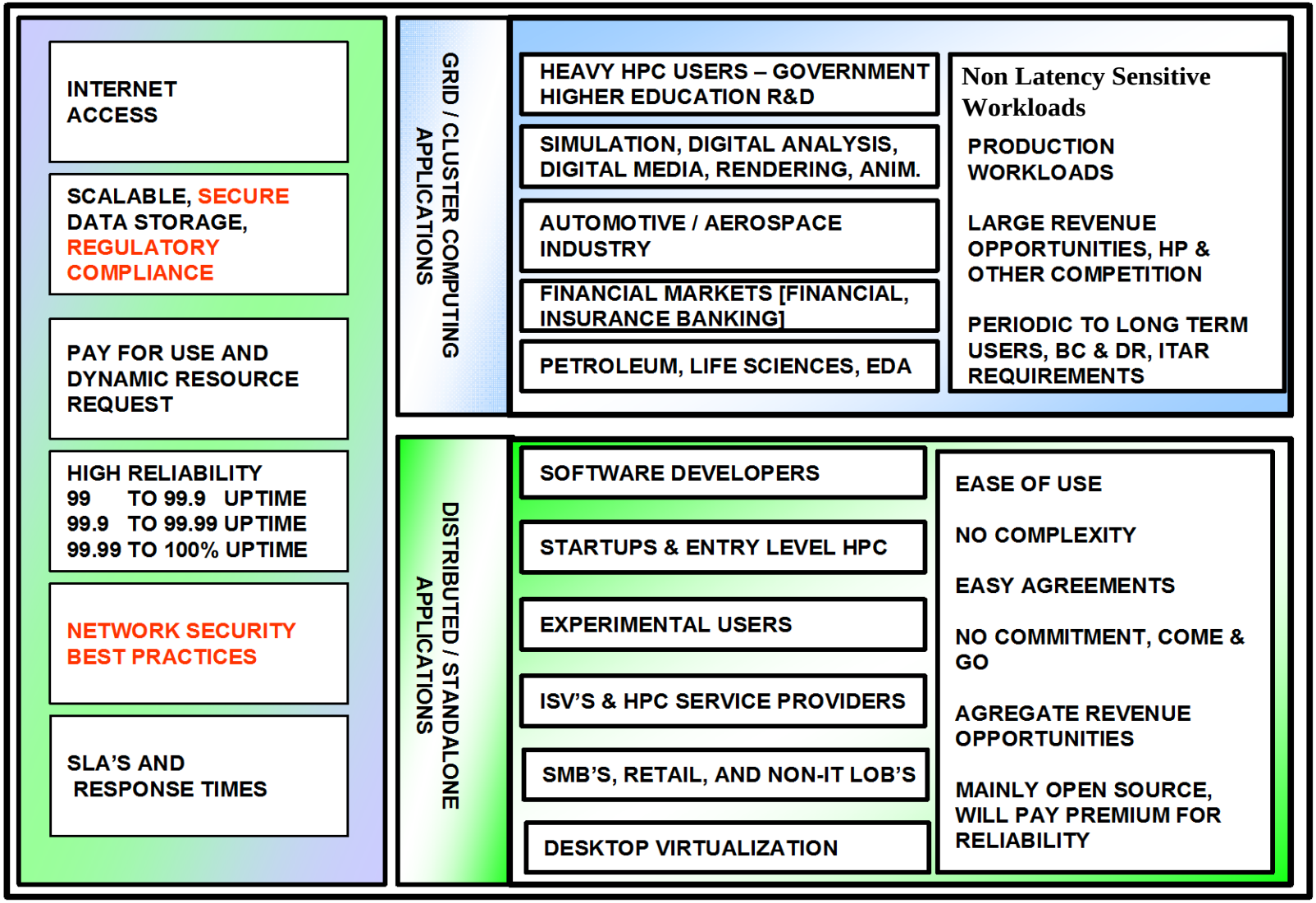
DCCoD



"More than 70% of the world's Global 1000 organizations will have to modify their data center facilities significantly during the next five years." - Gartner, September 2007

Energy costs 8x, Management costs 4x, 70% of IT budget is operational overhead

Cloud Computing spectrum – Observations



Srinivas Cheemalapati

Cloud support beneficial to HPC

- Application portability
- Image management
- Simulate scaling
- VM migration
- Better resource utilization
- Rapid provisioning

Application portability

- Users maintains their own virtualized OS
  - Linux, Windows, Solaris, ...
- Isolation:
  - No conflict: OS level, libraries, software, ...
  - Concurrent mix of different environments on same physical server
  - Subnetwork: cluster of servers

Image management

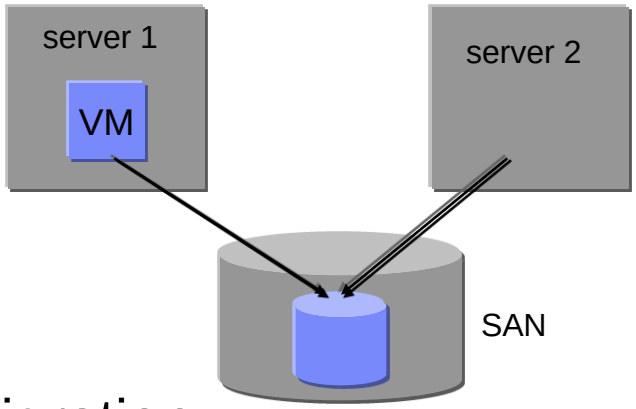
- Capture an HPC environment
  - Create & customize your own
  - Vendor supplied software stack
- Usage
  - Save, restore
  - Share, publish
  - Archive legacy environment
- Possible in any storage configuration
  - LVM, SAN LUN, VMWare

Simulate scaling

- Develop & debug large scale execution
  - Time consuming, system not productive
  - Result and execution time not a factor
- Shared cluster:
  - more efficient usage: don't need to tie up resource
  - when ready, execute on dedicated cluster

VM migration

- Fault tolerance
  - Detect pending failure: disk error, high temperature, communication error, ...
  - Migrate VM to healthy server
- Load balancing
  - High end servers
- Requirements
  - SAN, separate network for migration
  - Our observation: 300-400msec down time



Better resource utilization

- Consolidate multiple VM's on same node
- Multicore CPU in high end server
  - idle time due to disk I/O, blocking send/receive
- Network bandwidth
  - Applications have different communication requirement
- Tradeoff:
  - Flexibility in job time
  - More users, full system utilization

Rapid provisioning

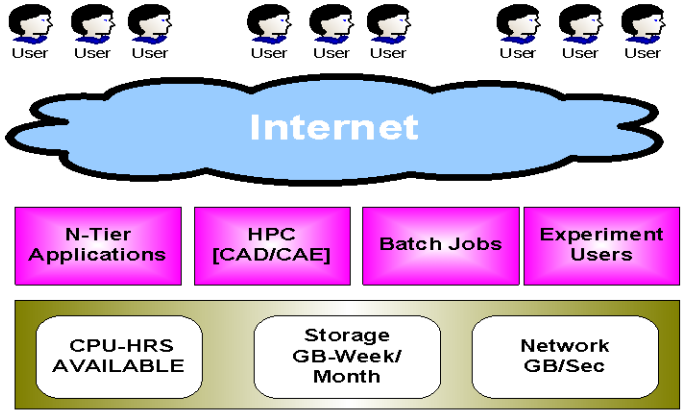
- New VM in about 2-3 minutes
- No work by IT team:
  - Cluster created automatically when requested, freed up automatically when done
- With self service portal:
  - Change thinking about IT infrastructure
  - Enable more experimentation
  - Vehicle for teaching

Concerns about Cloud

- Overhead
  - Virtualization cost
  - Computation, memory, disk I/O, networking
- Reliable performance
  - Resource sharing in VM's
  - Communication capacity

Building blocks for a cloud

=



Srinivas Cheemalapati
DCCoD

Not for distribution

The DCCoD models

Dedicated

- Custom Environment – choice of data center, servers, storage, dedicated switch, WAN/LAN High Availability Network
- Dedicated resources: 1-3 year commitment
- Leased by IBM Global Financing

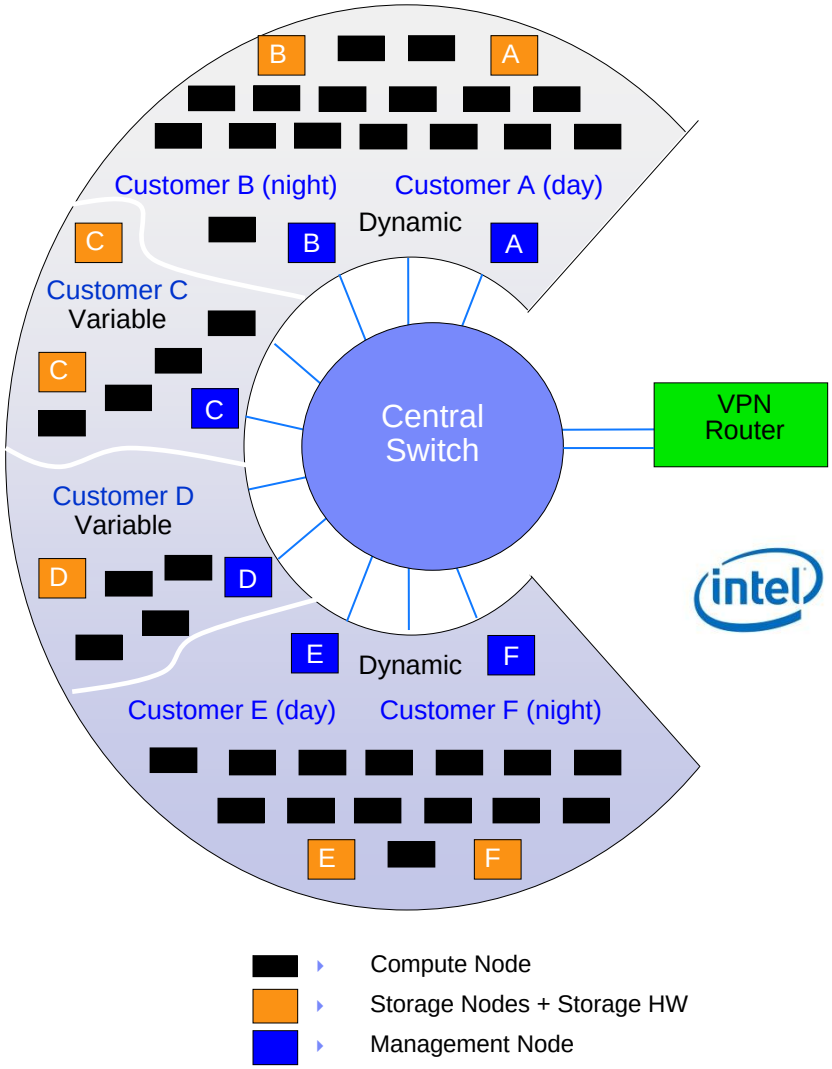
Variable

- Utilize DCCoD Center infrastructure
- Virtual clusters dedicated to one client at a time
- Automated network and server provisioning
- Reserve by the week; “Pay for use”

Dynamic

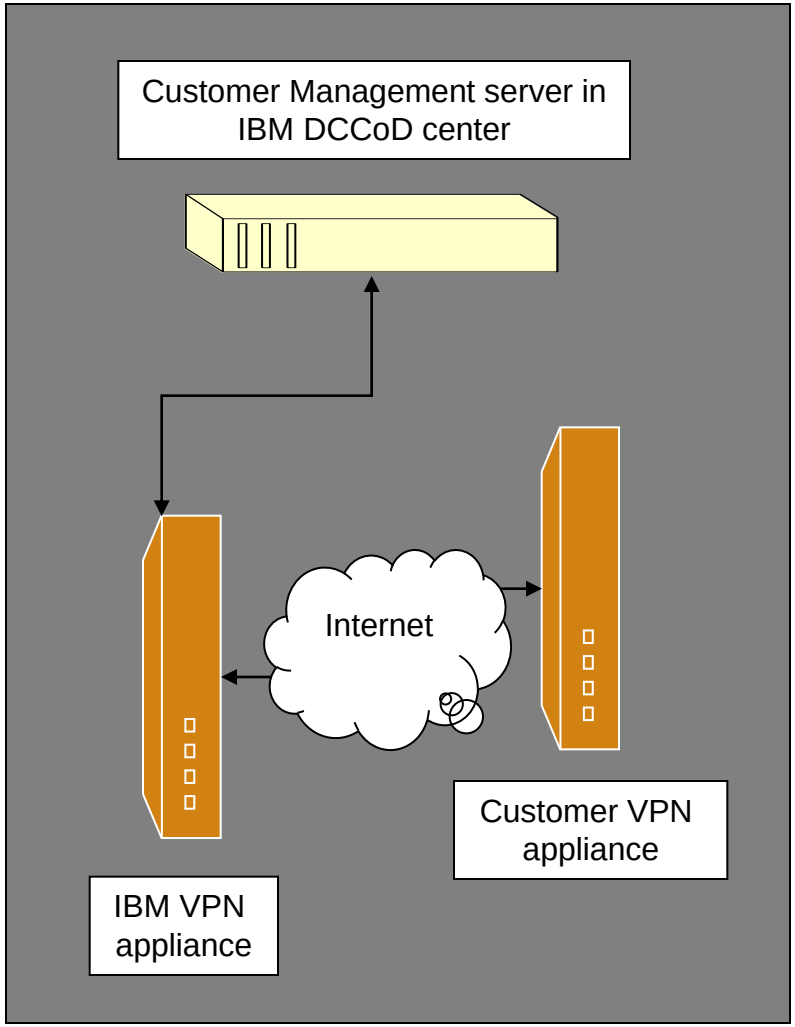
- Utilize DCCoD Center infrastructure
- Virtual clusters dedicated to one client at a time
- Automated network and server provisioning
- Reserve by the hour*; “Pay for use”
- Ideal for Intraday and post trading workloads

*Minimum commit: Monday – Friday, 8 - 10 hrs per day

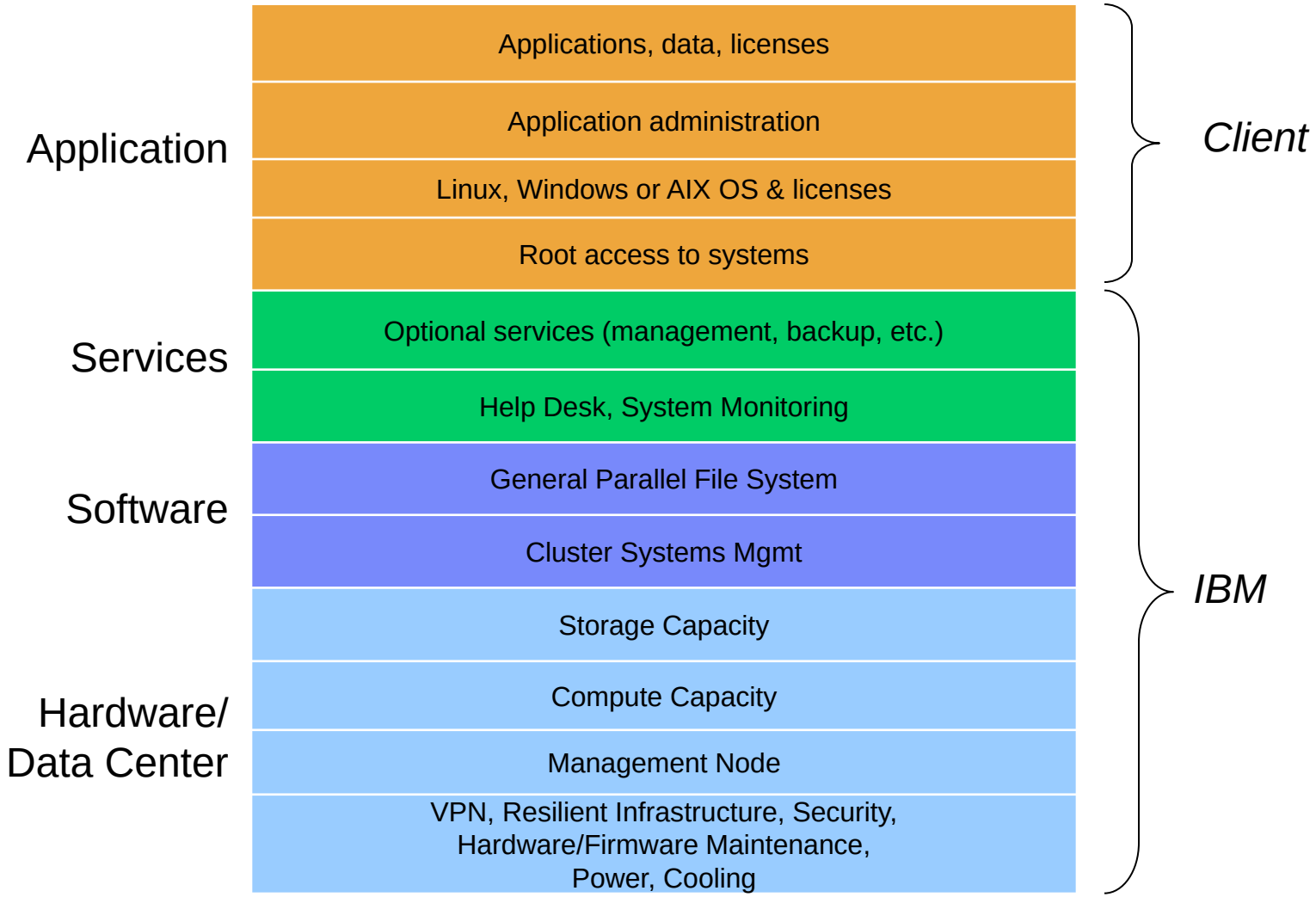


IBM DCCoD base access

- Renewable “annual membership”
- Proactive advanced planning to support on demand usage
- Ongoing Virtual Private Network (VPN) access
  - Software standard
  - Various hardware options available
- Ongoing management server footprint for test and engagement readiness
- Choice of three models to contract for compute capacity
  - dedicated, variable, dynamic



IBM & client responsibilities



Examples of Cloud Computing: "Services as a Service"

1. **Software as a Service (SaaS)**
   - Applications are delivered through the browser to multitudes of customers.
   - Customers avoid upfront investment in servers and software licensing.
   - Providers spread application hosting costs over large customer base.
2. **IT Infrastructure as a Service (Utility Computing)**
   - Data center infrastructure elements (storage, virtual servers, and even virtualized data center partitions) are provided on demand over the Web.
   - Initial adoption is mainly for supplemental, non-mission-critical needs, but broader use may follow.
3. **Web Services as a Service**
   - Web service providers offer APIs that enable developers to exploit functionality over the Internet, rather than delivering full-blown applications.
4. **Development Environments as a Service**
   - Vendors provide development environments as a service.
   - Clients build applications that run on the provider's infrastructure and are delivered to end users via the Web from the provider's servers.
5. **Managed IT Services as a Service**
   - Services such as a virus scanning service for e-mail or an application monitoring service are provided as managed services to IT groups rather than to end-users

See http://www.infoworld.com/article/08/04/07/15FE-cloud-computing-reality_1.html for more details

IBM HPC MANAGEMENT SUITE FOR CLOUD

Overview

IBM HPC Management Suite for Cloud offering allows customers running technical computing and analysis workloads to consolidate their scattered cluster infrastructure, increase hardware utilization, gain access to larger cluster infrastructure and deploy their HPC applications in a Cloud environment.

Key Client Needs Addressed include

- Consolidation and efficient sharing of HPC infrastructure
- Ease of manageability and access to HPC infrastructure via a self-service web portal
- Automation (Rapid provisioning, cluster set-up, power & energy management) capabilities to optimize usage of compute resources and increase system utilization
- Centralized user management, usage metering & accounting.

Target Markets

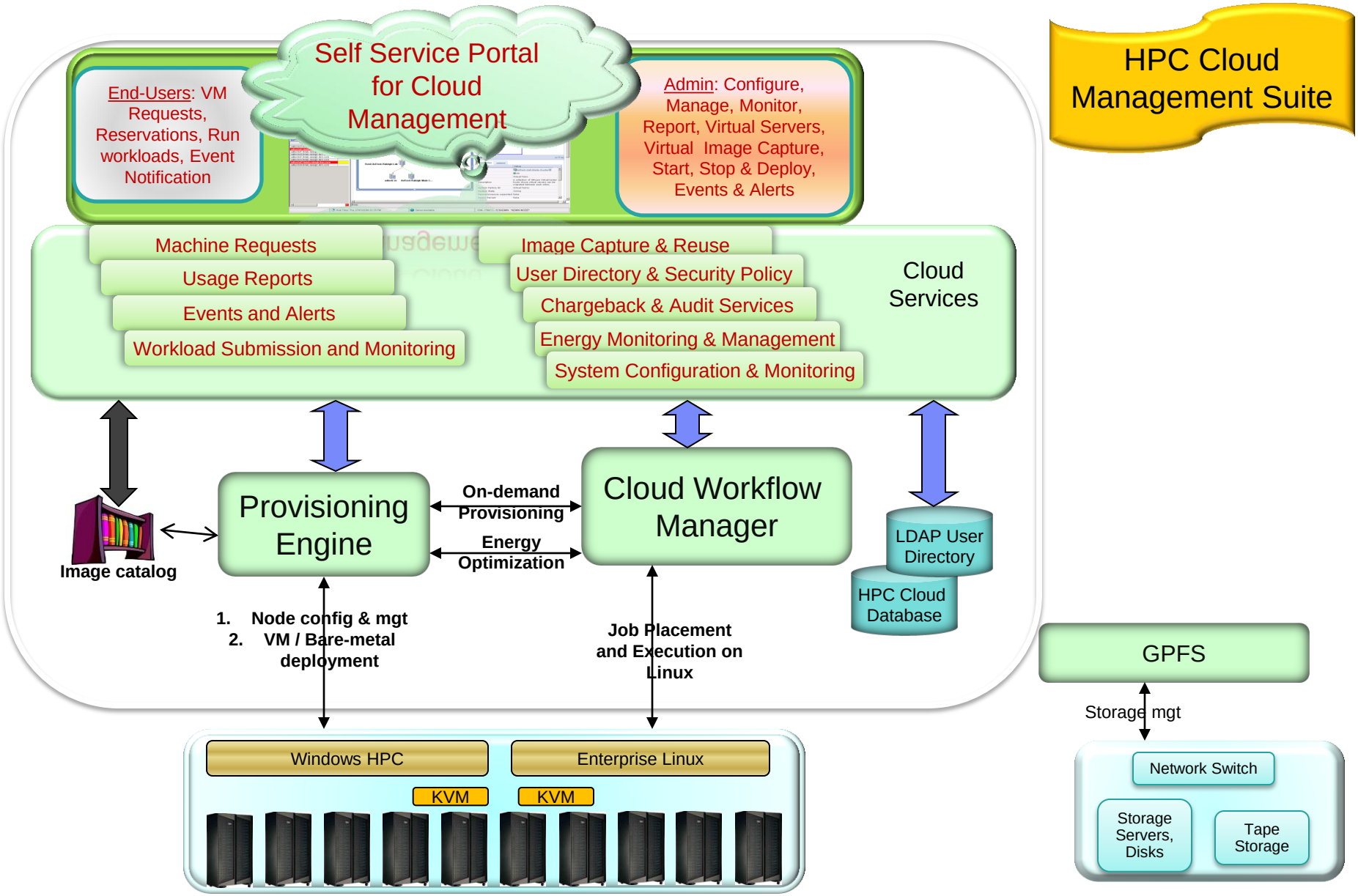
High Performance Computing / Technical Computing and Analysis workloads in Industrial sector – Automotive, Aerospace & Defense, Chemical & Petroleum, Electronics, Financial Services, Government, Research institutes, Higher education and Life Sciences

State-of-the-art Cloud Services for High Performance Computing

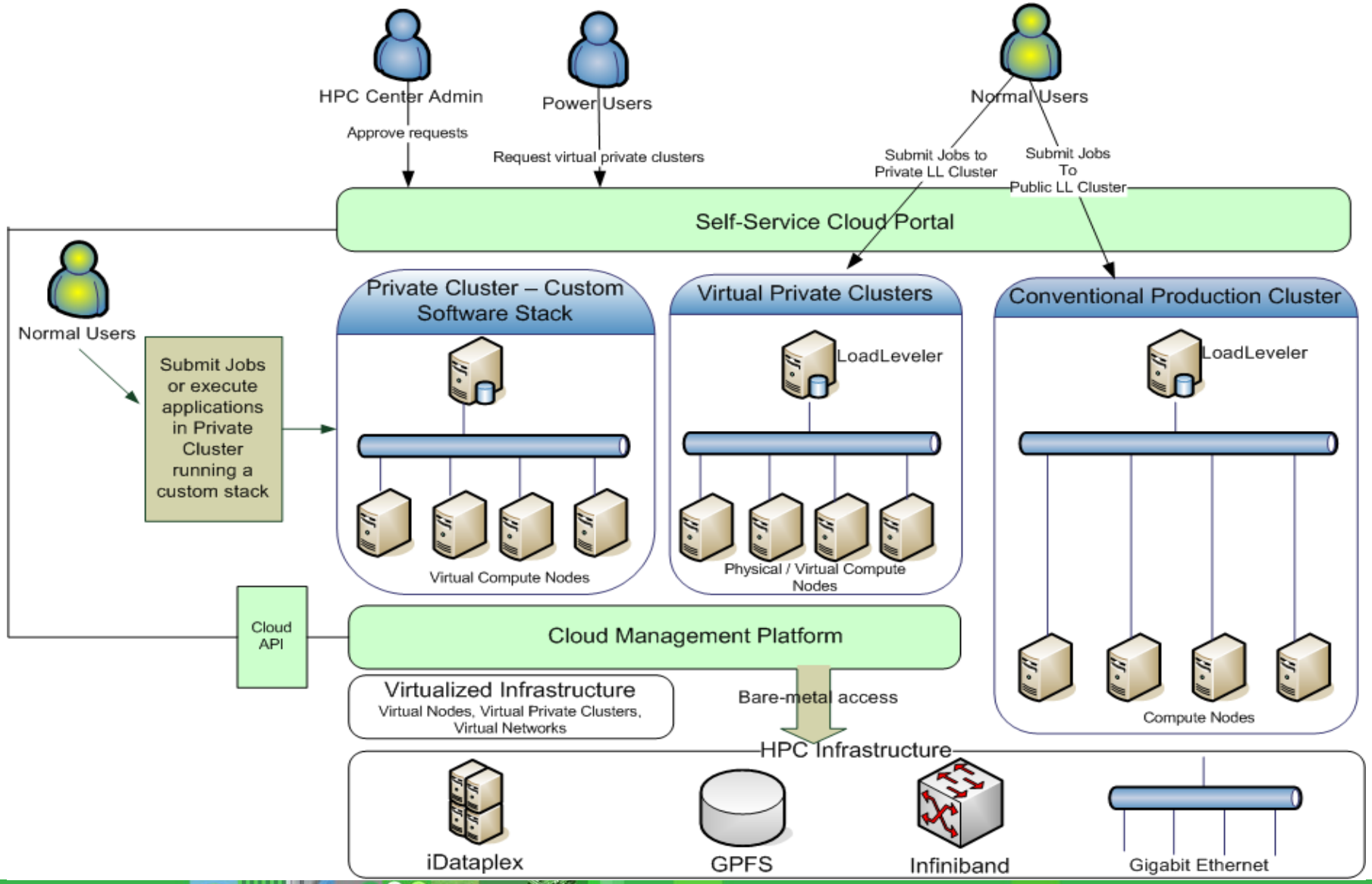
- Self service web portal
  - Common Cloud services interface for users and administrators
  - Cloud Management Policy Administration
  - Job submission and monitoring multiple job queues
- Rapid image deployment
  - Diskless, Bare-metal and Virtual Machine provisioning
  - Deploy Linux and Windows images
  - Image capture and reuse
- Power management
  - Policies for power state management with ability for manual overrides from portal
  - Energy monitoring and usage records for provisioned nodes

State-of-the-art Cloud Services for High Performance Computing

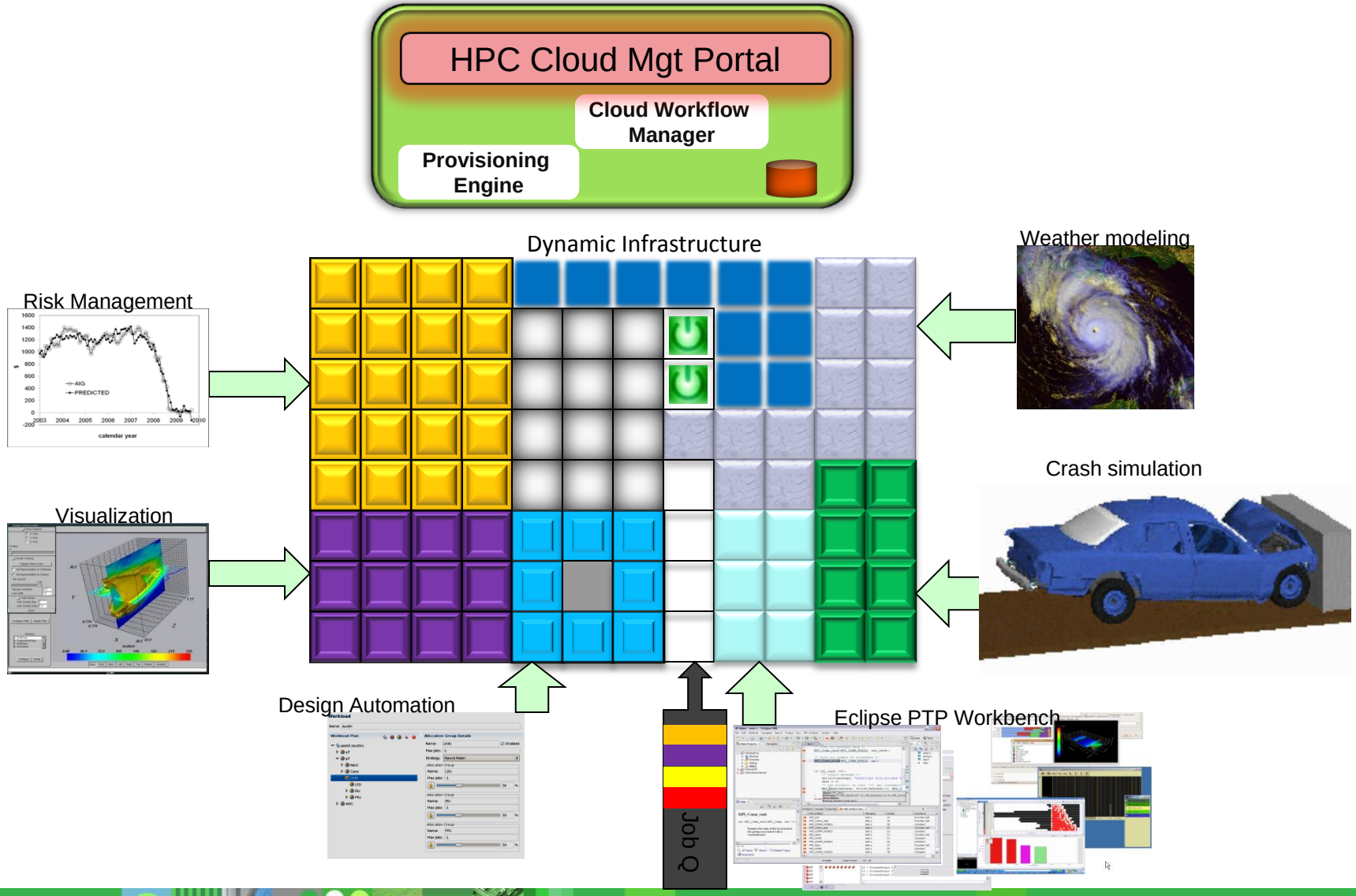
- Automated setup and configuration for HPC workloads
  - Configuration options allow high resource utilization
  - Multiple clusters (out-of-box support with LoadLeveler) within a private cloud
  - Single-point submission and monitoring for multiple LL job queues
- Network partitioning
  - Virtual LAN setup for private clusters
- Usage Metering & Accounting
  - Compute resource usage metrics captured in database
  - Ability to create customized billing reports
- Centralized user management and security
  - LDAP based authentication



USE CASES FOR RELEASE 1.0

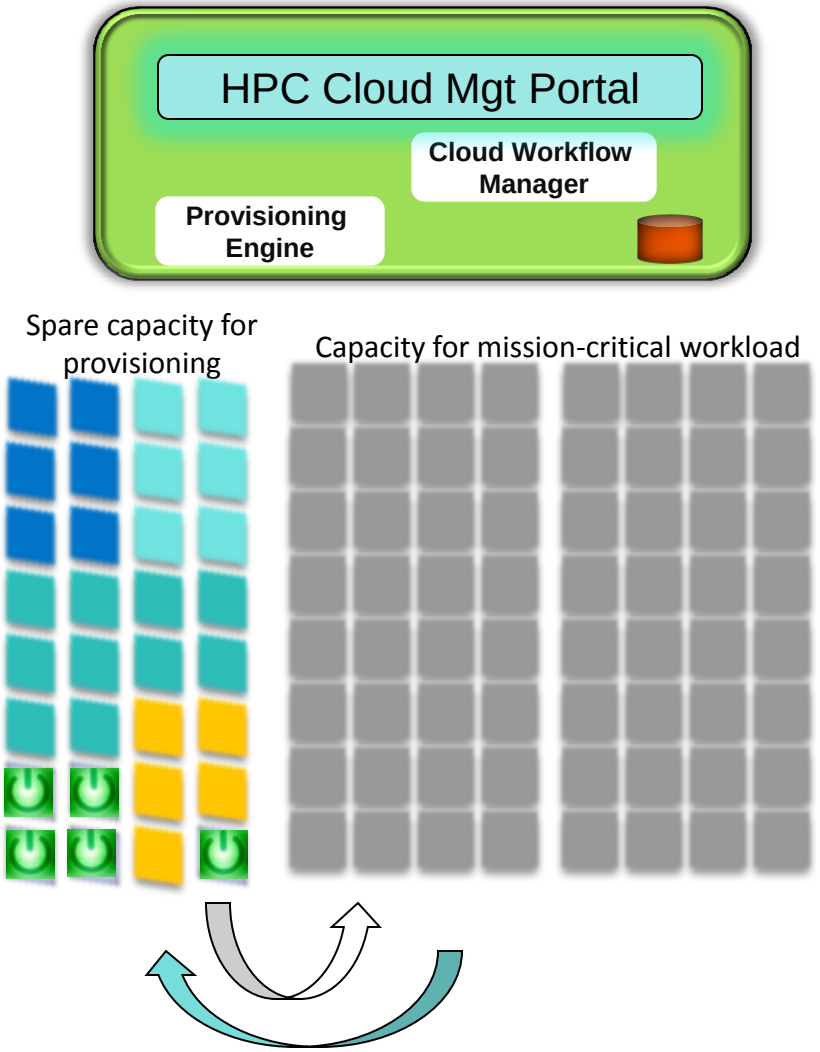


Orchestrate clusters for HPC applications...



... while running mission-critical workload

- All machines in the cloud are able to run mission critical workload when desired
- Spare capacity is configured to run low priority work
- A provisioning request from a HPC Cloud, the workload will be vacated and the nodes will be provisioned.
- Provisioned nodes are restored to default image for running normal workload before rejoining
- Idle nodes go into power save or power off mode



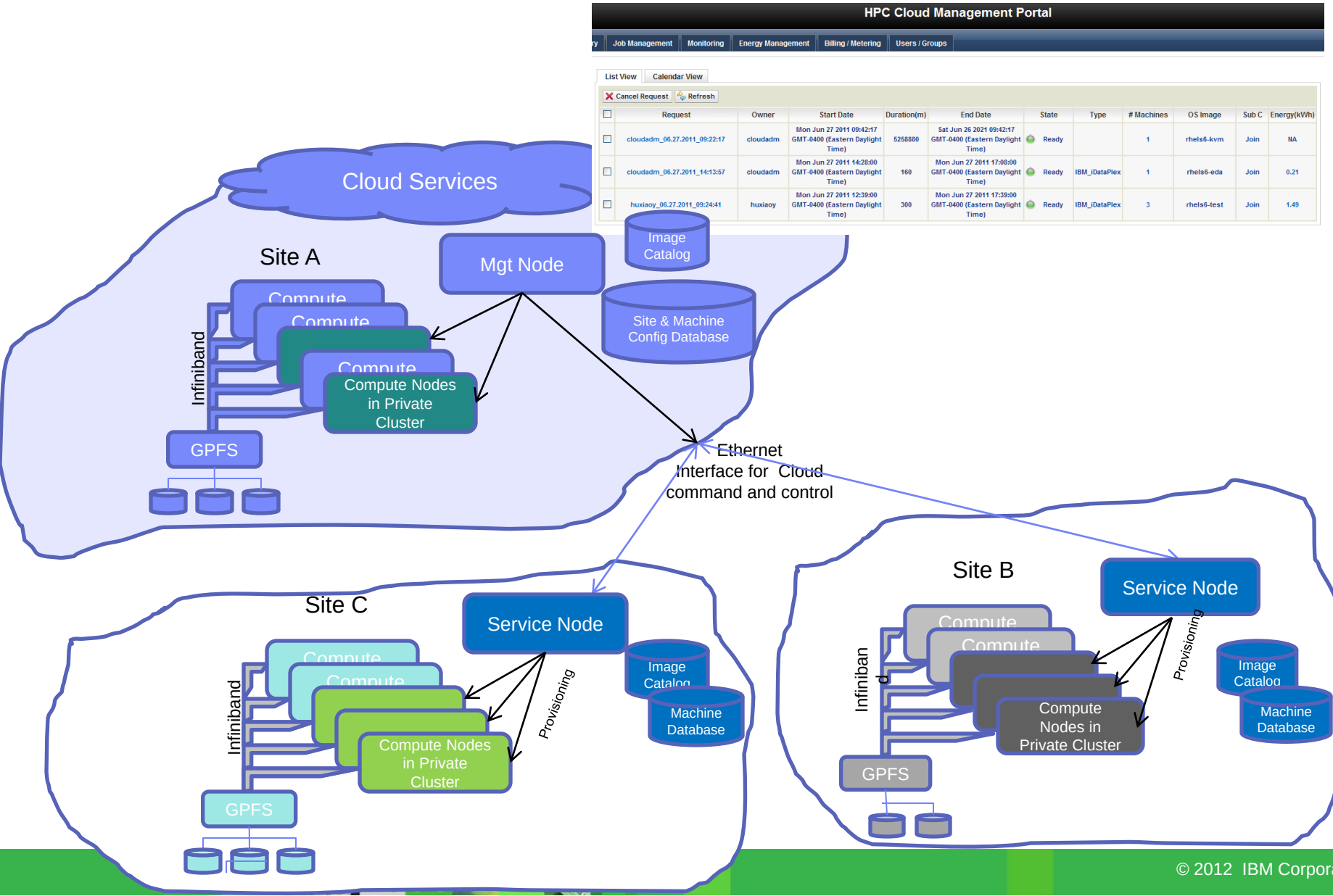
In Summary, benefits of an optimized HPC cloud management suite

- **Visibility**
  - monitor your HPC Cloud on the web

- **Control**
  - keep your HPC Cloud in its desired state using policy-based administration
- **Automation**
  - manage huge and growing infrastructures while controlling cost and quality

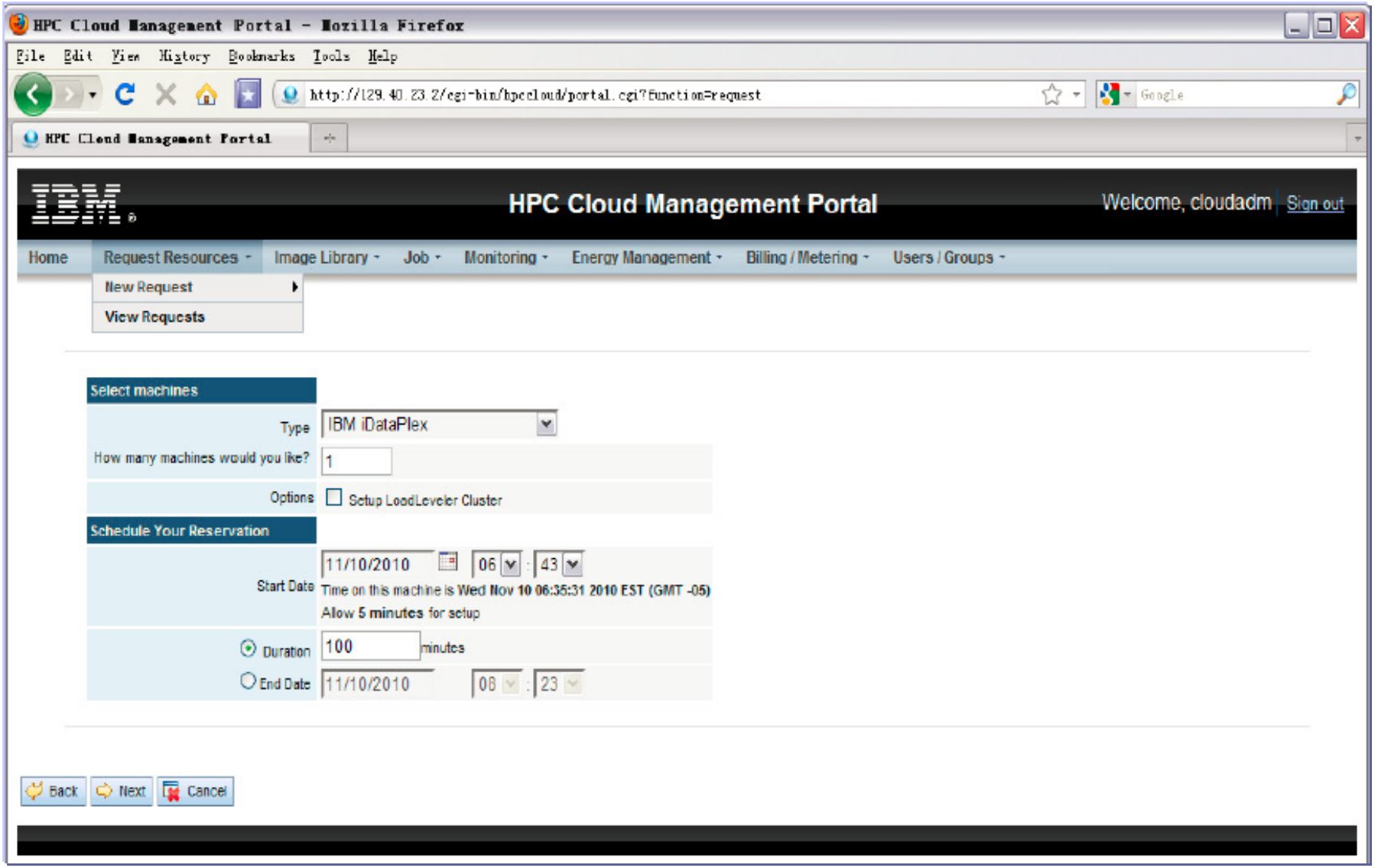
- **Performance**
  - experience the full performance and scalability of a proven HPC platform

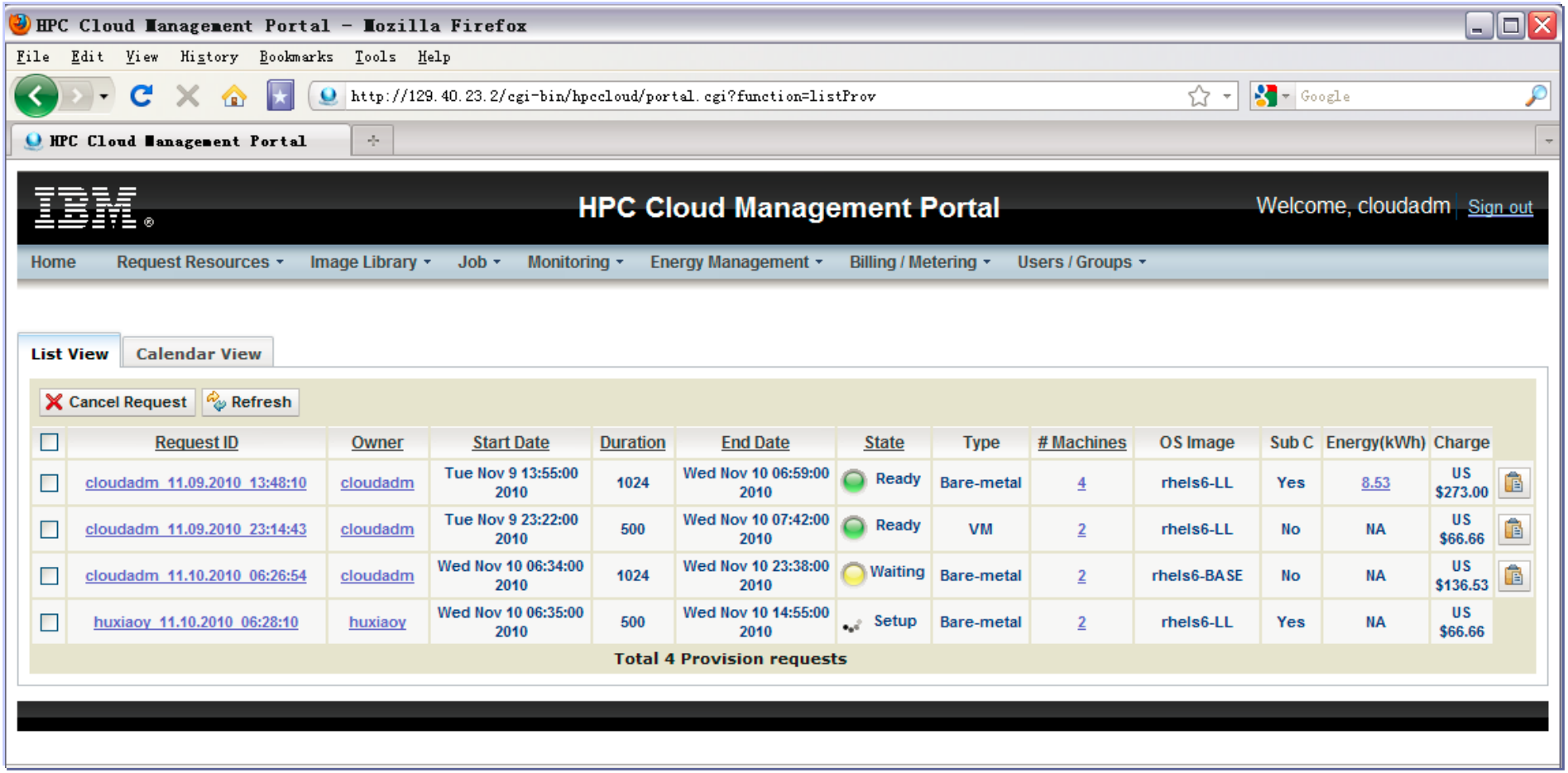
Multi-site HPC cloud vision for large enterprises

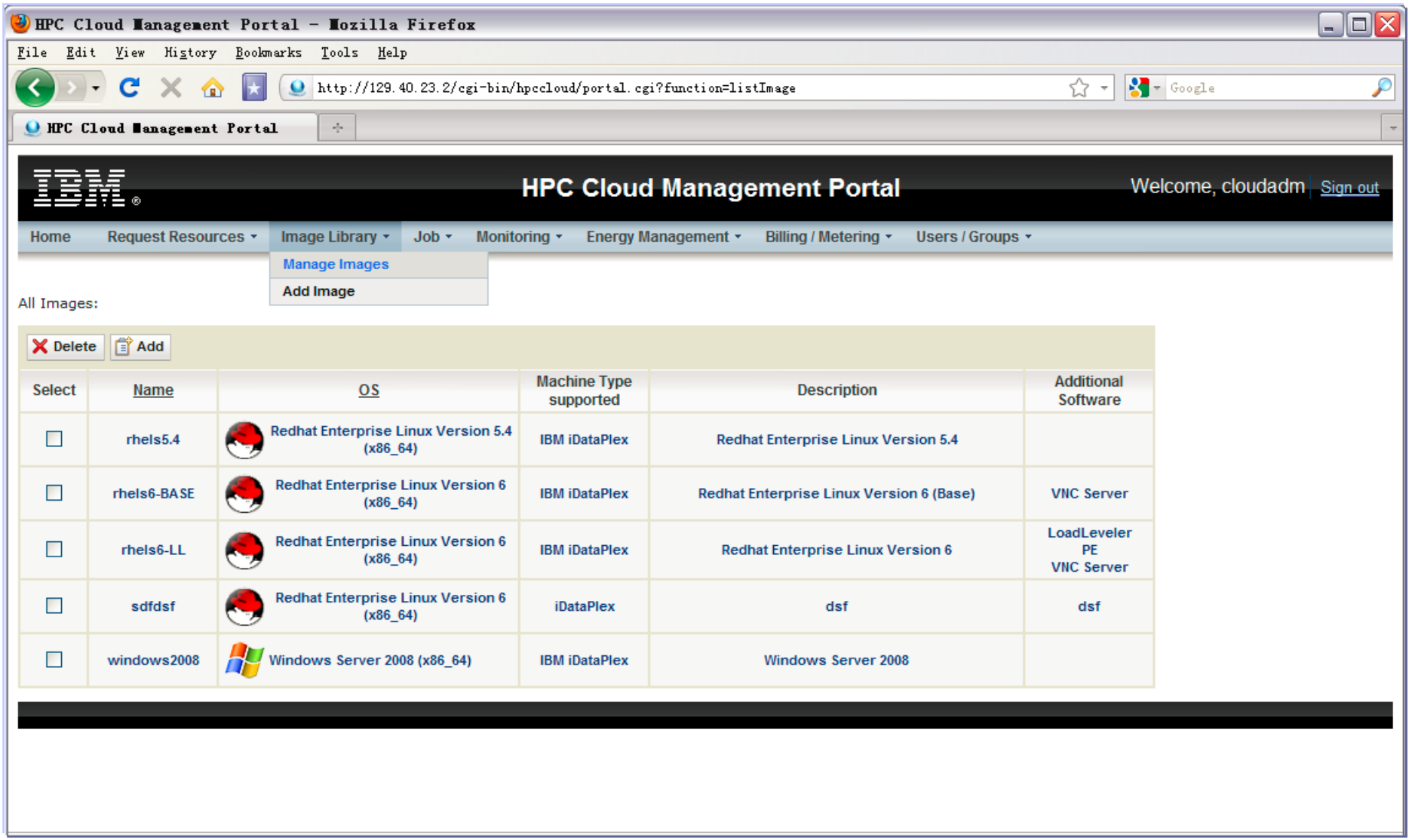


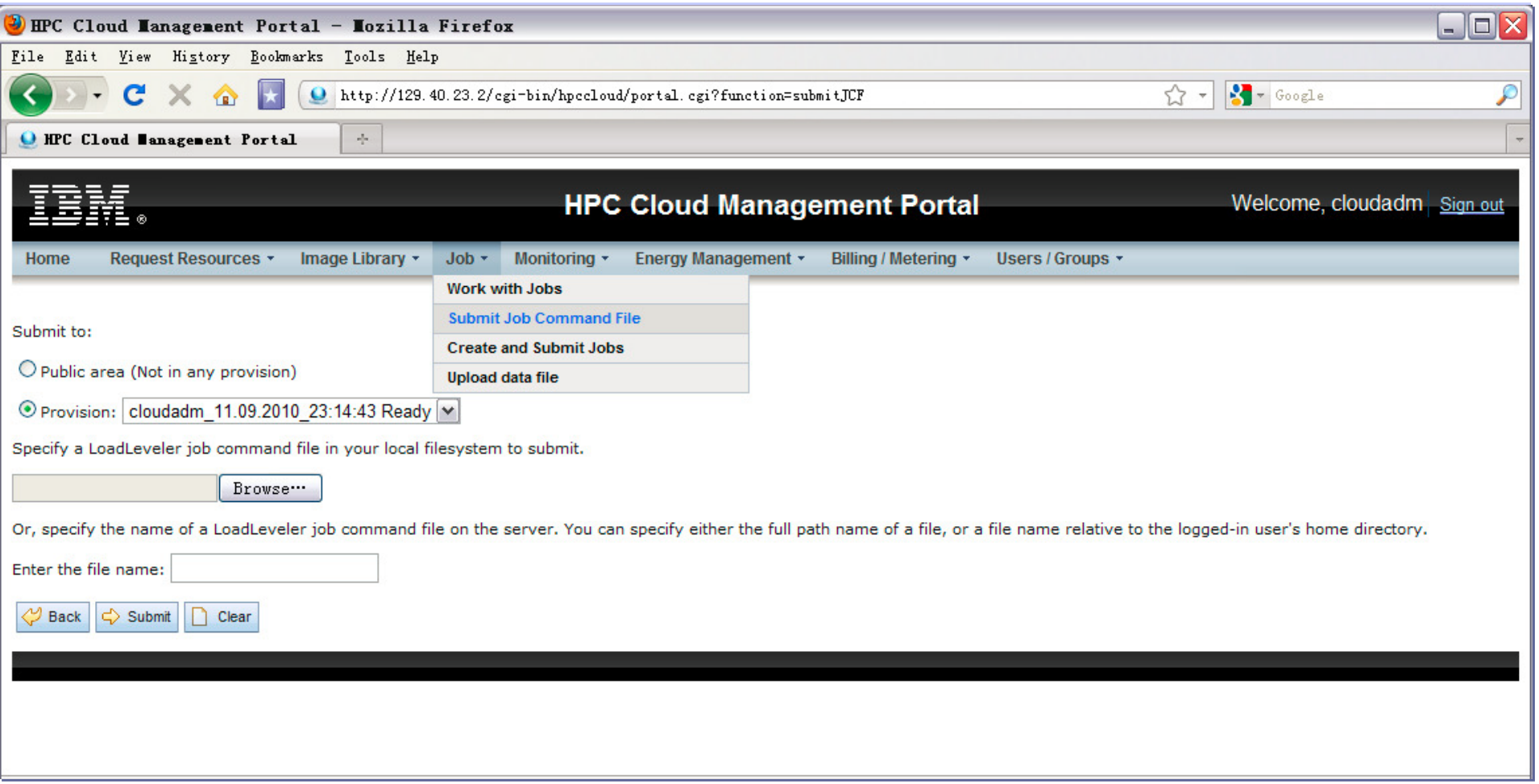
Example of:

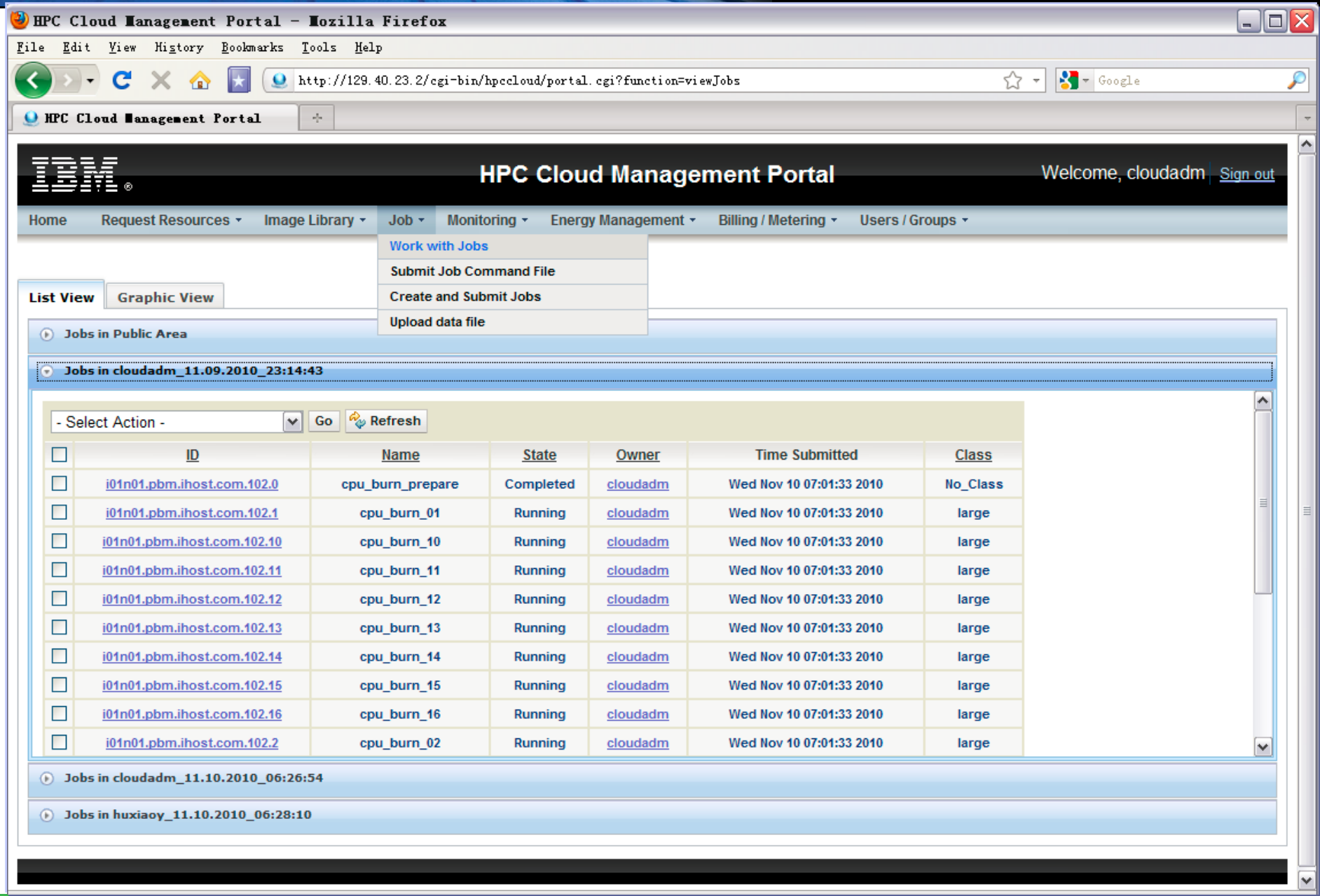
IBM HPC CLOUD PORTAL

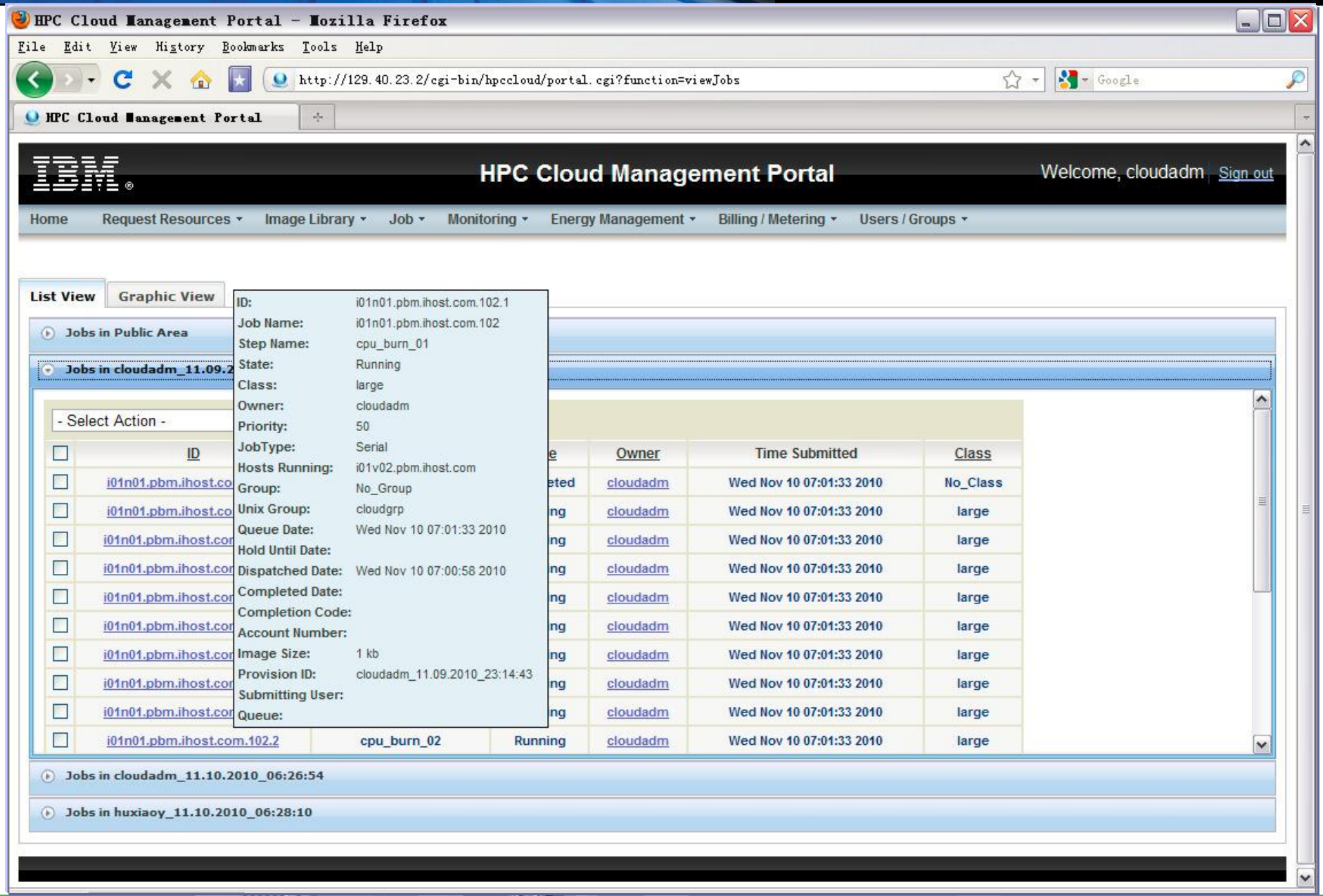


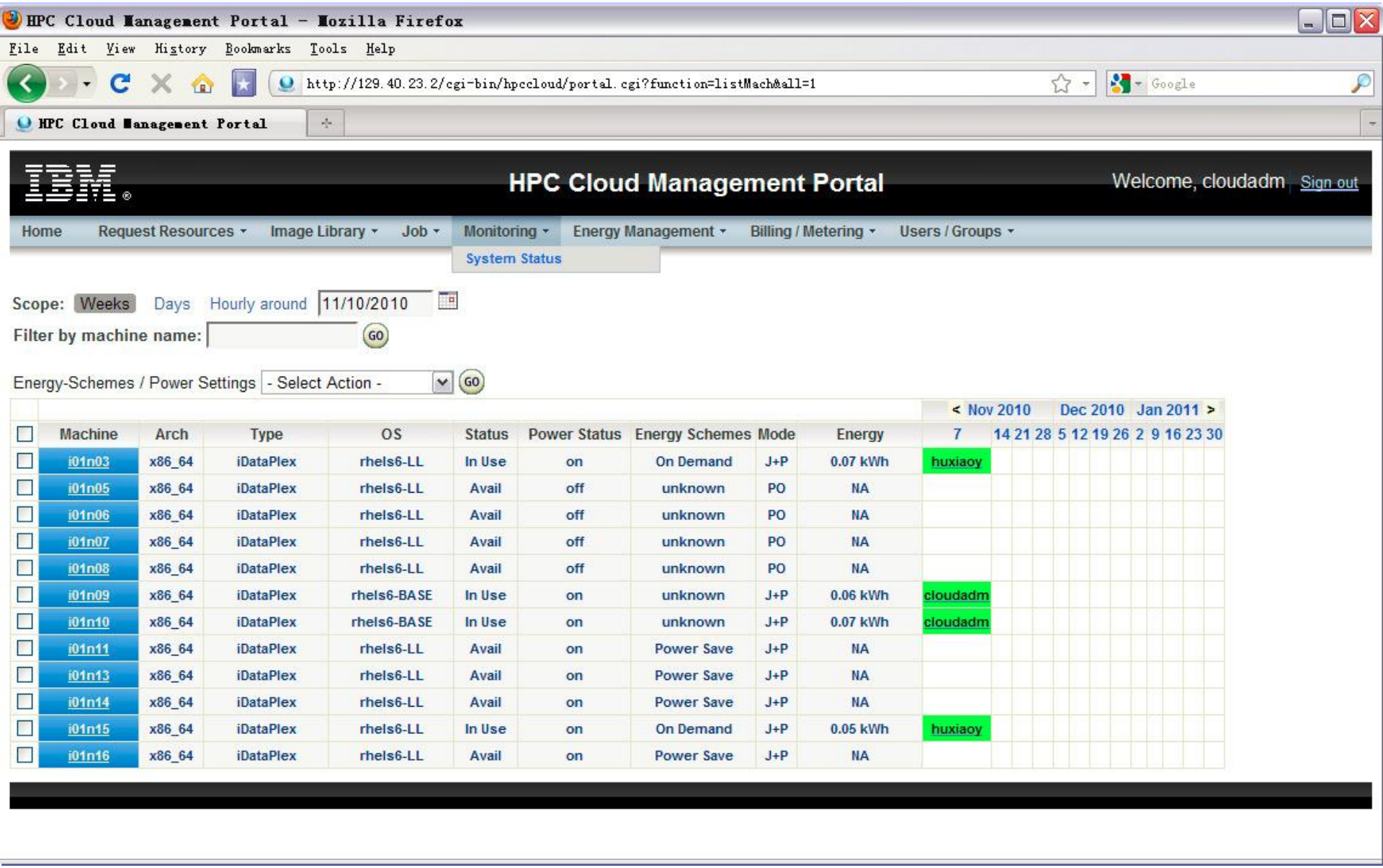


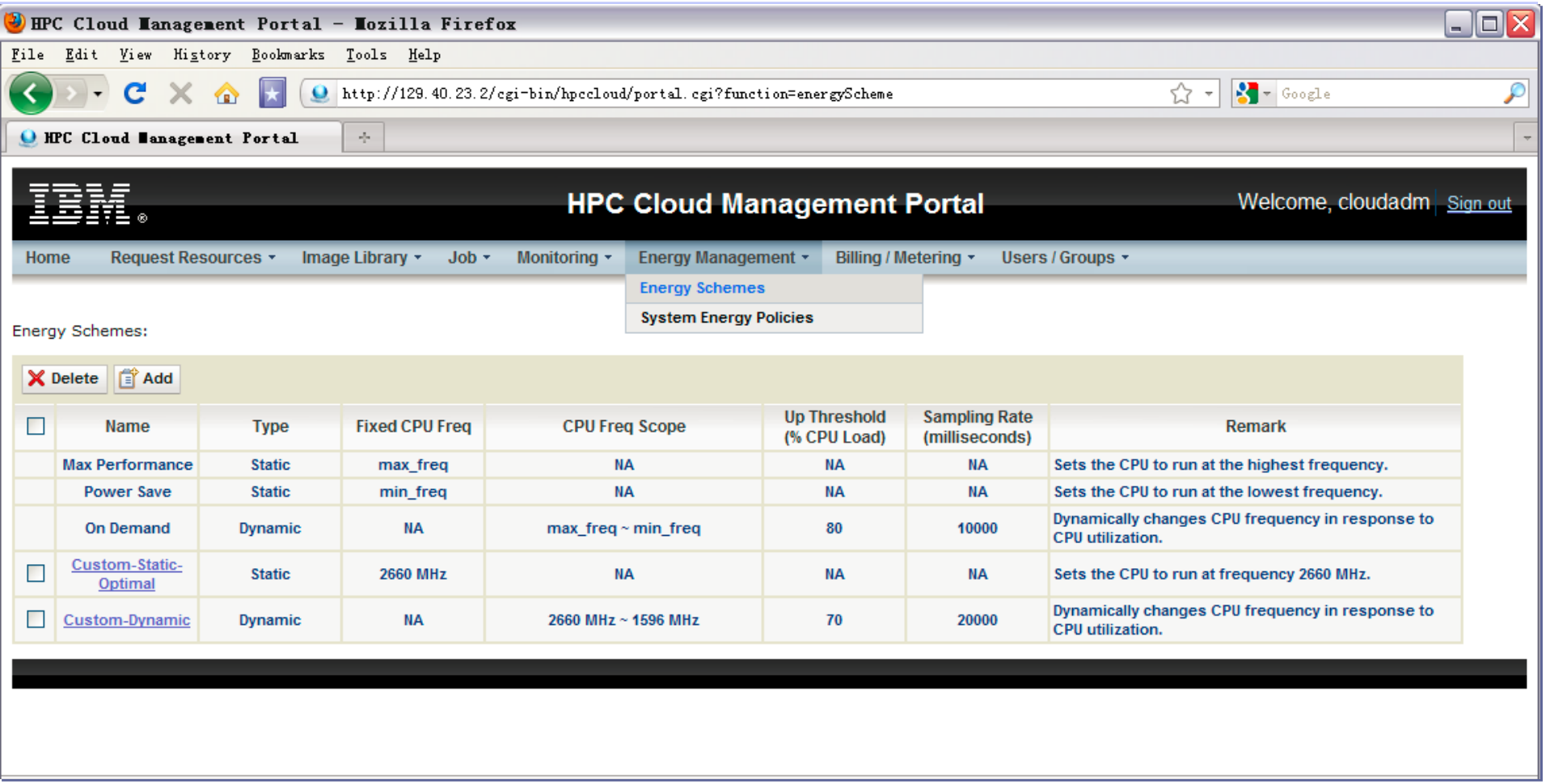




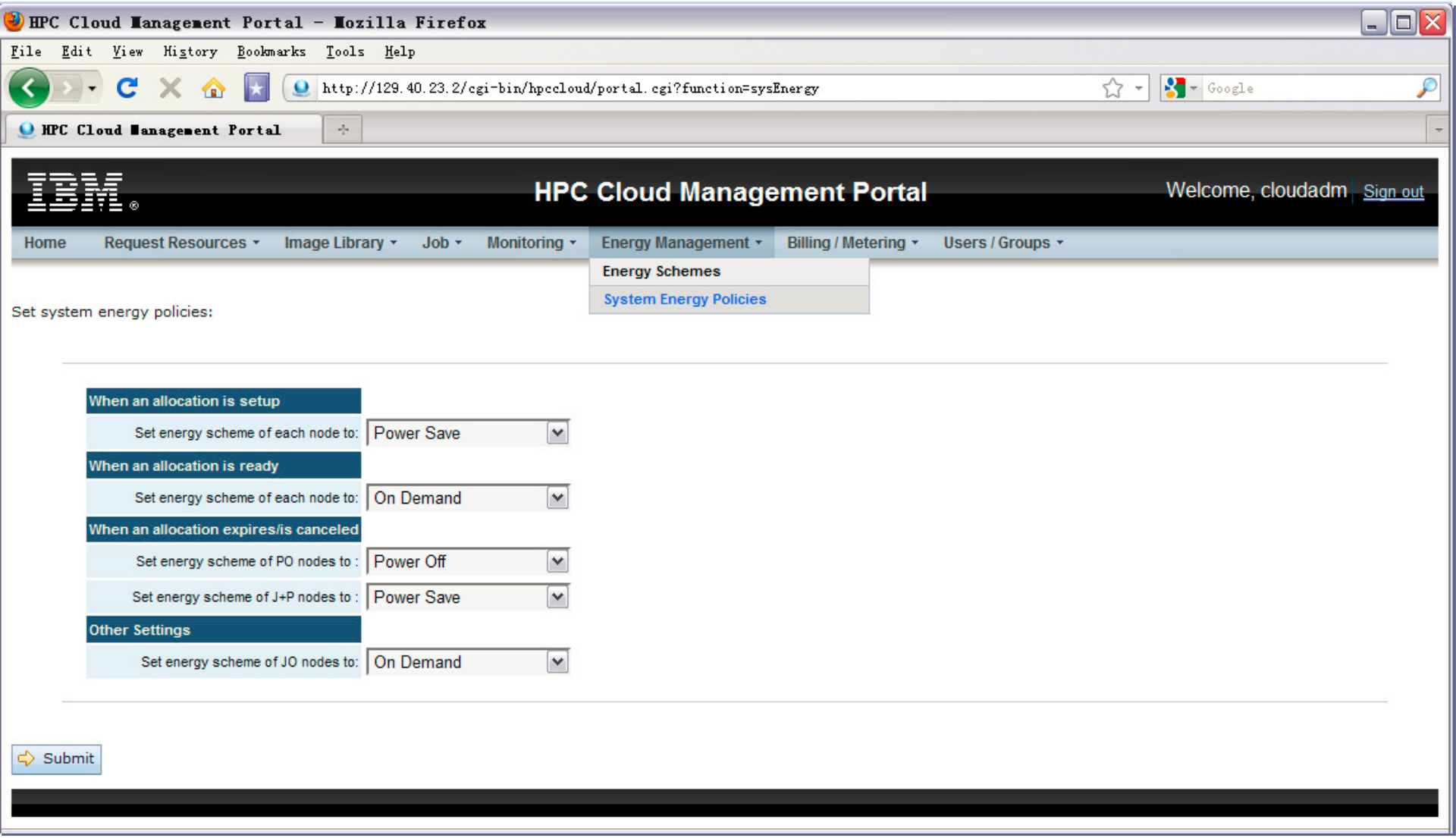


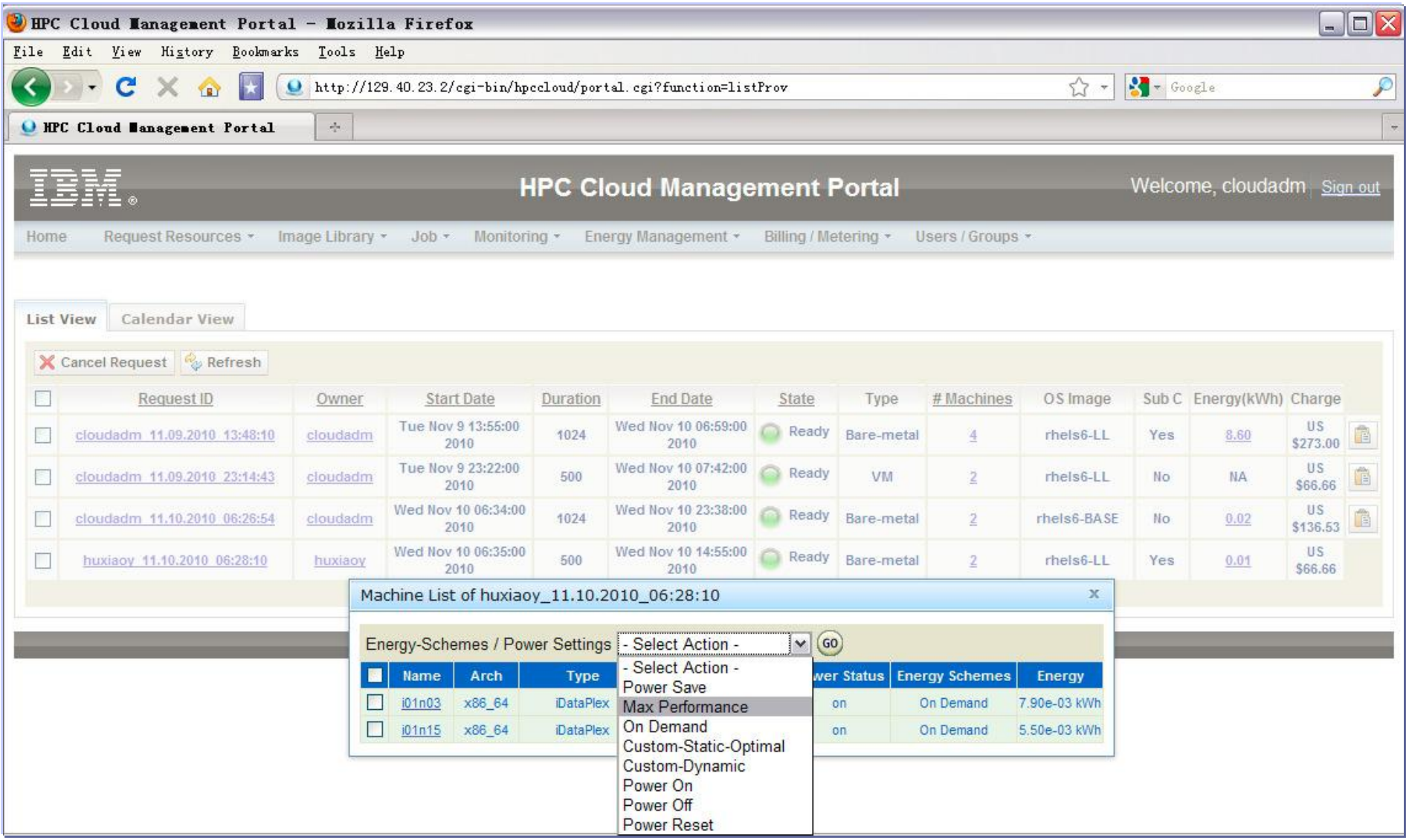


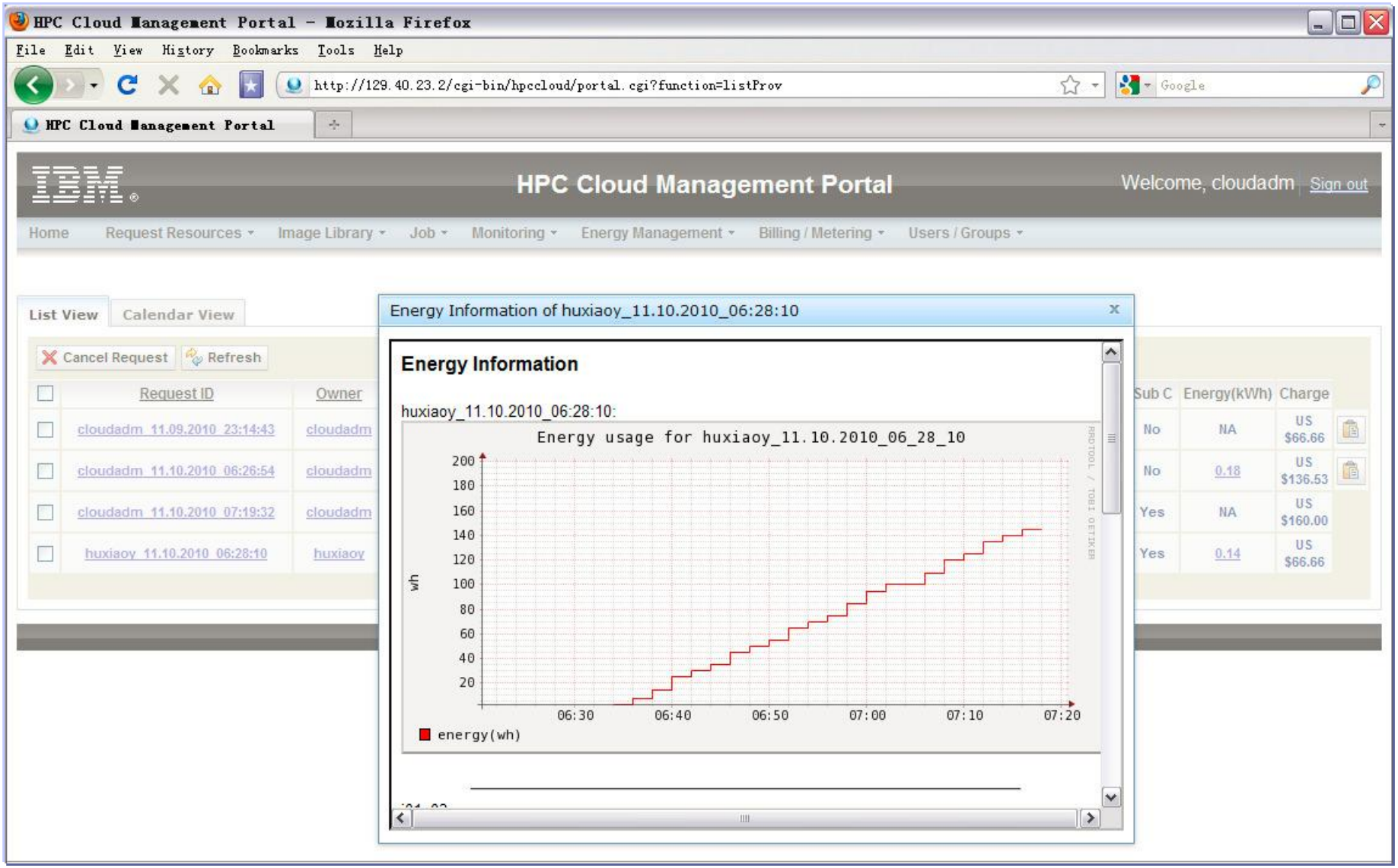


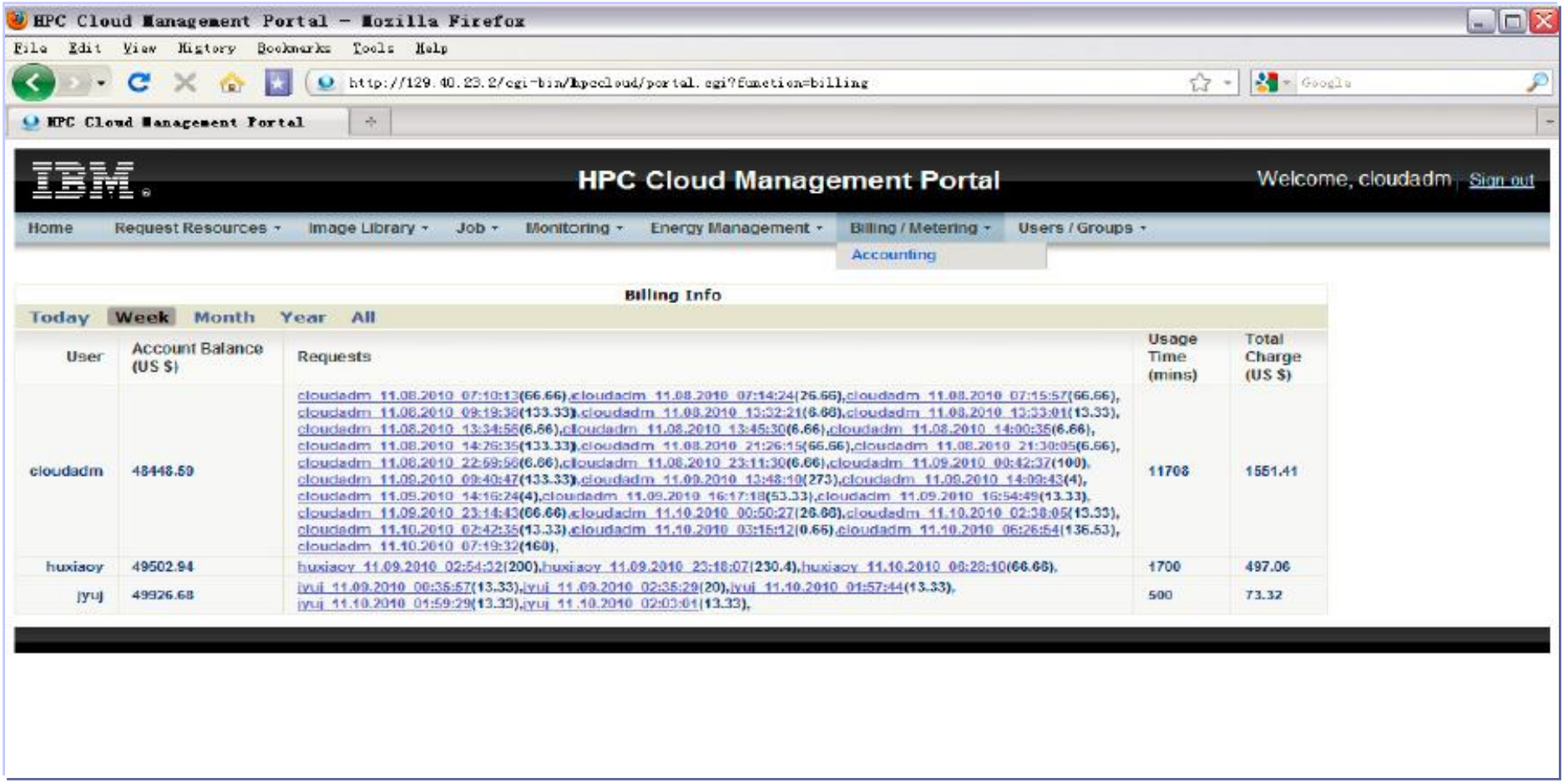


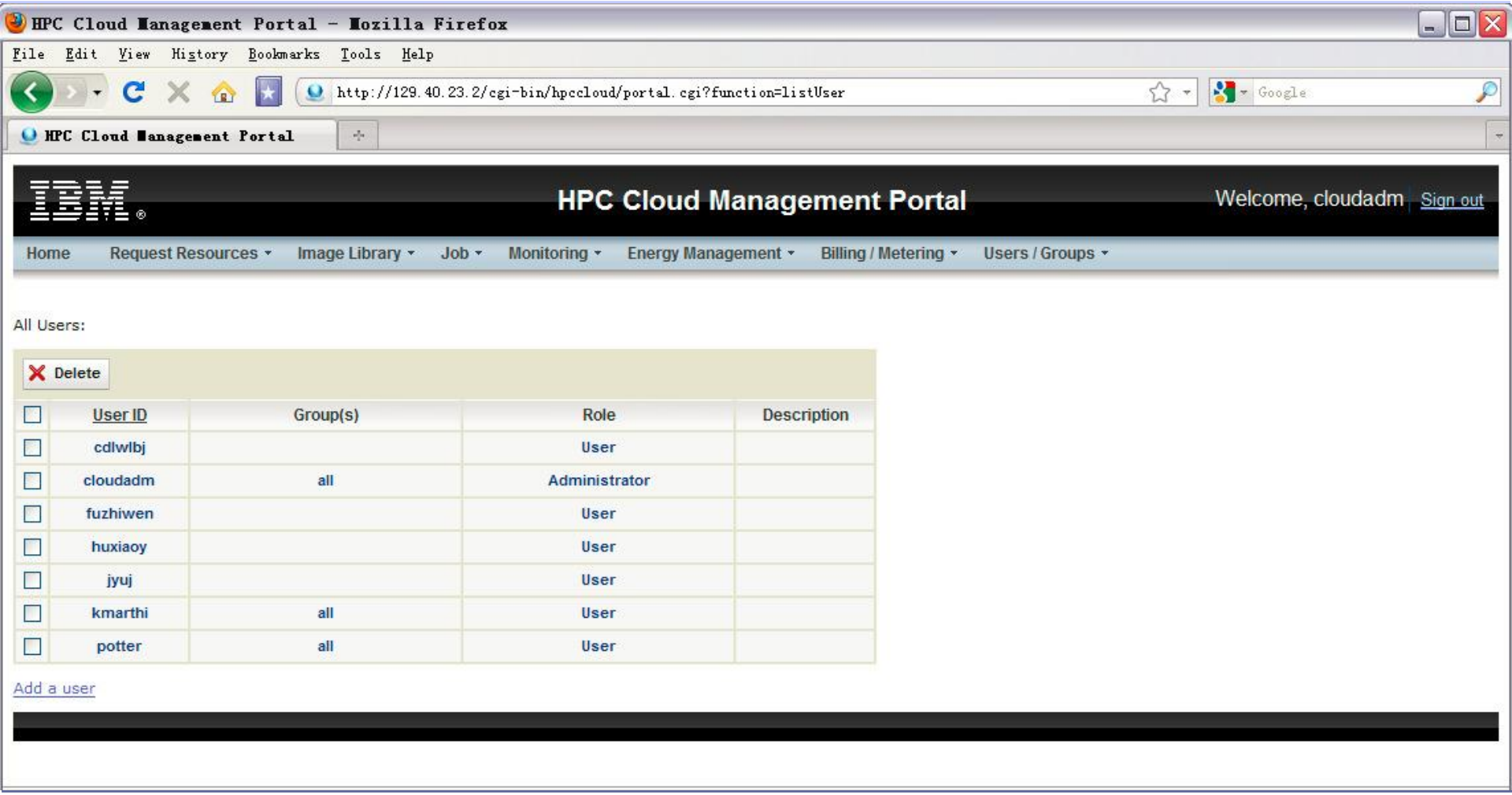
	Name	Type	Fixed CPU Freq	CPU Freq Scope	Up Threshold (% CPU Load)	Sampling Rate (milliseconds)	Remark
	Max Performance	Static	max_freq	NA	NA	NA	Sets the CPU to run at the highest frequency.
	Power Save	Static	min_freq	NA	NA	NA	Sets the CPU to run at the lowest frequency.
	On Demand	Dynamic	NA	max_freq ~ min_freq	80	10000	Dynamically changes CPU frequency in response to CPU utilization.
	Custom-Static-Optimal	Static	2660 MHz	NA	NA	NA	Sets the CPU to run at frequency 2660 MHz.
	Custom-Dynamic	Dynamic	NA	2660 MHz ~ 1596 MHz	70	20000	Dynamically changes CPU frequency in response to CPU utilization.











IBM WORLD COMMUNITY GRID

Every time Henrik Erikson
takes a break,
AIDS research advances.

Here's a smart way of benefiting humanity.
Use your computer's downtime to add power to World Community Grid.
Join today at www.worldcommunitygrid.org

A3

William Wong is
hard at work
for cancer research.

Here's a smart way of benefiting humanity.
Use your computer's downtime to add power to World Community Grid.
Join today at www.worldcommunitygrid.org

A3

world community grid
technology solving problems

HOME | ABOUT US | RESEARCH | FORUMS | STATISTICS | MY GRID | DOWNLOAD | Select Language | HELP

Register > Select Projects > Download > Explore

Member Name: florin.manaila

- Select Projects
- Download software
- Install software
- Explore World Community Grid

Check the box(es) for the project(s) to which you want to contribute your spare computing time. Click continue when you are finished. Don't worry, you can make different selections later if you change your mind.

Available Projects

- [ ] Participate in All Projects
- [x] Say No to Schistosoma
- [x] GO Fight Against Malaria
- [x] Drug Search for Leishmaniasis
- [ ] The Clean Energy Project - Phase 2
  - Please review the system requirements before opting to participate in this project
- [x] Help Cure Muscular Dystrophy - Phase 2
- [x] Help Fight Childhood Cancer
- [x] Help Conquer Cancer
- [x] Human Proteome Folding - Phase 2
- [x] FightAIDS@Home

Please note: Some projects may occasionally have periods where no work is available. If you choose to receive work for only one project, the project may, at times, not have work units available for download. For this reason, we recommend that you select more than one project.

continue

Visit the Forums

If you need technical assistance beyond what's available in Help, please visit the forums to post your questions and get answers on how others are using World Community Grid.

Related Resources

- Software Overview
- System Requirements
- Points System
- Member Policy

POWERED BY IBM

world community grid.
technology solving problems

HOME | ABOUT US | RESEARCH | FORUMS | STATISTICS | MY GRID | DOWNLOAD | Select Language | HELP

Register | Select Projects | Download | Explore

Member Name: florin.manaila

- Projects selected: Human Proteome Folding - Phase 2

- Download software
- Install software

- Explore World Community Grid

The World Community Grid software will be downloaded for:

Windows ®
(6.10.58)
(7, Vista, XP, Me, 2000)
md5: c12b490e917b500b1c664d7c43acdc35

If you'd like to download the software for a different operating system, click here.

When the download box appears, click "Save," and choose a location (Desktop is a good choice). When the download is finished go to where you saved the file and double-click to start the installer.

• Click the button to start the download: download now

NOTE: When you have successfully installed the software and connected your computer to World Community Grid, you will be taken to the next page automatically. If you think that you have completed the process correctly, but the page does not advance, then please look at the following items to help you determine what additional steps may need to be taken:

- Print a full set of instructions by clicking here
- Look here to see if your firewall might be stopping you
- You may need to configure a proxy

If these items do not help, then please consider getting help from other members in the Support Forums. Or if you would rather, you can use our contact us page to ask for assistance.

Visit the Forums

If you need technical assistance beyond what's available in Help, please visit the forums to post your questions and get answers on how others are using World Community Grid.

Related Resources

- Software Overview
- System Requirements
- Points System
- Member Policy

POWERED BY IBM

florin.manaila@ro.ibm.com

THANK YOU