

Current Topics in High Performance Computing (HPC): Clouds, RAS, and True Cost of Ownership (TCO)

The mainstreaming and growth of HPC has driven the industry to examine approaches to increase reliability, reduce TCO, and explore cloud computing.



October, 2009

Dr. Srinu Chari
chari@cabotpartners.com
Phone: 203 205 0705

<http://www.cabotpartners.com>

Prepared for SURA IT Meeting 2009
In Washington DC

***Cabot
Partners***

Outline



- ❑ High Performance Computing (HPC) Recent Applications Trends
- ❑ HPC Data center challenges and solutions
- ❑ Cloud computing for Commercial and HPC enterprises
- ❑ Workload classification-making the case for cloud computing
- ❑ A workload based cost-benefit framework for IaaS Cloud solutions using the Uptime Institute model for TCO
- ❑ HPC reliability, availability, and serviceability (RAS)

Established Mainstream Cross-Industry HPC Applications

3

- ❑ **Academic and Research Computing including weather/climate prediction**
 - traditional lead users
- ❑ **Financial Trading**
 - low latency algorithmic trading applications in an explosive market data environment
- ❑ **Life Sciences**
 - *in silico* biology, high throughput screening, computational chemistry, imaging, and systems biology for drug discovery and development
- ❑ **Engineering**
 - virtual product design, structural analysis, fluid dynamics, and circuit simulation.
 - smaller suppliers are embracing HPC
- ❑ **Upstream Petroleum**
 - seismic analysis and reservoir simulation

Emerging Cross-Industry HPC Applications



- ❑ **Media and Entertainment**
 - rendering, pattern recognition, interactive games
- ❑ **Search, Security, and Surveillance**
 - video surveillance, border control, biometrics
- ❑ **Business Intelligence/Analytics**
 - retail, financial services, and health care
 - getting insight from growing volumes and diversity of operational data

Data center challenges

- ❑ Escalating energy costs – about 11% annually while new server spend expected to be flat¹
- ❑ Industry is defining new metrics for energy efficiency²
- ❑ Data centers account for about 25% of enterprise IT budgets and is growing at 20% annually³
- ❑ Investment required to build a large-enterprise data center has risen to \$500M, from \$150M, over the past 5 years and larger data centers take 2 years or more to design and build and are expected to last for 12 years³
- ❑ Server utilization typically tops off at 5%-15% and results in wasted energy and unemployed capital³
- ❑ Difficult to forecast whether a 50% increase in demand would require 25% or 100% more server and data center capacity³
- ❑ HPC reliability, availability, serviceability concerns growing
- ❑ True cost of ownership (TCO) and ROI must be carefully examined

1. Jed Scaramella, "Worldwide Server Power and Cooling Expense 2006-2010 Forecast", September 2006.

2. John R. Stanley, Kenneth Brill, and Jonathan Koomey, "Four Metrics Define Data Center "Greenness"", White Paper, Uptime Institute.

3. McKinsey on Business Technology, Innovations in IT Management, Number 14, Winter 2008.

Data center solution approaches

- ❑ **McKinsey³ suggests a centralized governance model with the CIO empowered by the CEO to manage data centers by:**
 - managing IT assets aggressively through virtualization,
 - providing incentives to IT personnel to improve forecasting and minimize deviations from real demand,
 - treating data center resources as scarce resources and ensuring that business units implement a true cost of ownership (TCO) model for new systems and applications,
 - implementing new metrics for data center efficiencies that account for energy, utilization, and floor space.

- ❑ **What's also needed is a comprehensive approach especially for HPC and emerging web and analytic workloads that includes:**
 - energy-efficient and computationally dense systems,
 - software for better systems and power management and utilization,
 - Improved reliability, availability, and serviceability (RAS), and
 - flexible delivery models such as cloud computing – Infrastructure as a Service (IaaS) that adapt easily to computing demands.

3. McKinsey on Business Technology, Innovations in IT Management, Number 14, Winter 2008.

What is the promise of cloud computing?



- ❑ Provides dynamically scalable and often virtualized IT (hardware, software, and applications) resources as a service transparently over the Internet to users who need not have knowledge of, expertise in, or control over the technology infrastructure.
- ❑ Incorporates software as a servicee (SaaS), Web 2.0 and other popular, recent, Internet computing trends, and also builds upon recent IT infrastructure solution concepts such as grid computing, utility computing, and autonomic computing.
- ❑ Smaller organizations may no longer need large capital outlays in hardware or facilities to deploy their services or the labor to operate these IT facilities.
- ❑ Various business and deployment models are emerging
- ❑ Some workloads are more amenable to clouds than others – highly variable and very compute intensive benefit most initially

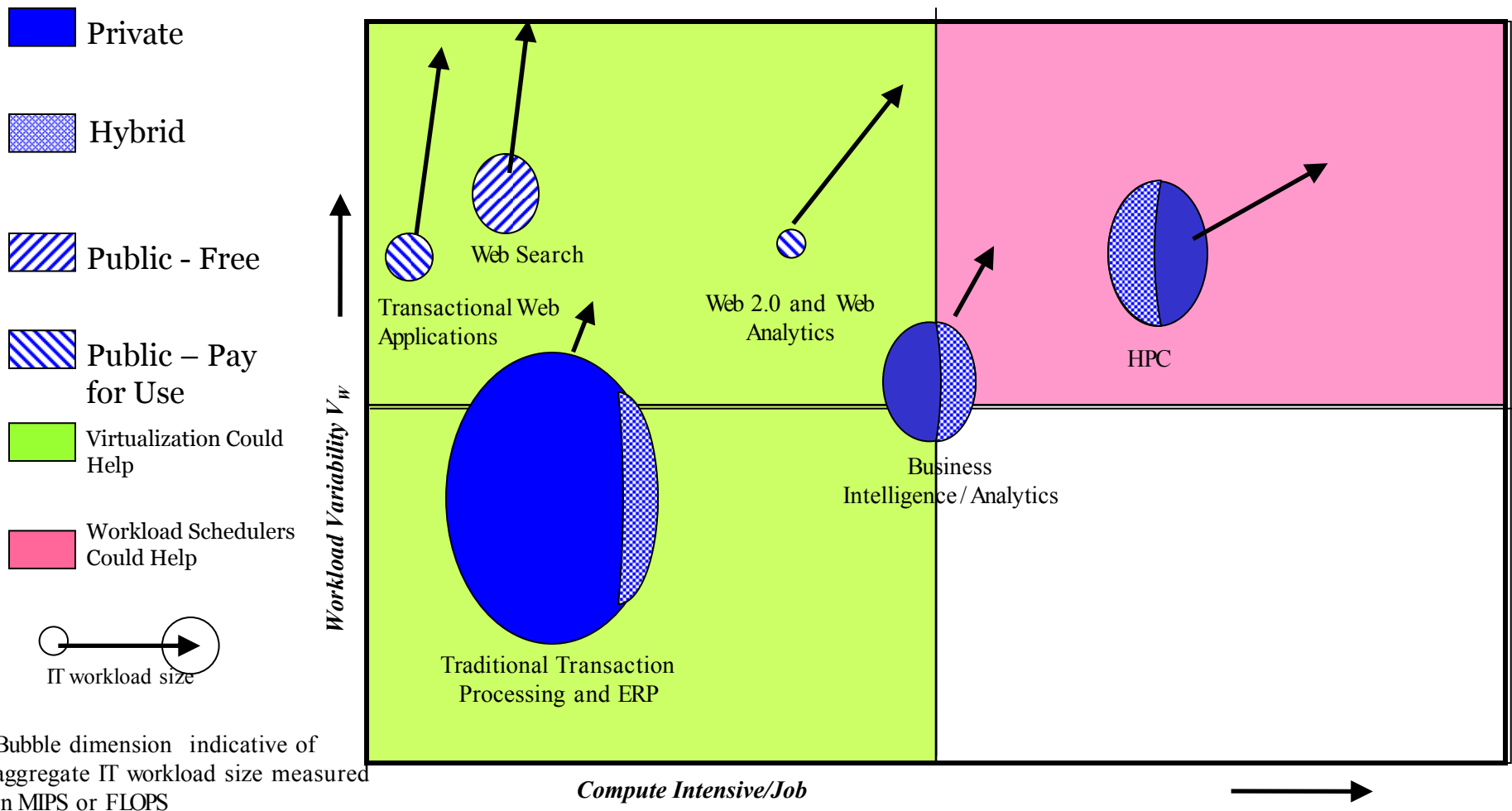
IT Clouds vs. Airline Transportation – Transport IT to the Cloud!

Unlike typical analogies of cloud computing to commodity electric utilities, IT clouds are more analogous to the airline transportation industry with a range of business models.

IT Industry	Transportation	Cost Model	Business Impact
In-House Data Center or Private Cloud (not-outsourced)	Own planes	CapEx, OpEx, long term	Total control and responsible for entire infrastructure including service levels
Private Cloud (outsourced)	Lease planes and service	CapEx and OpEx in medium term lease	Control of infrastructure but not fully responsible for service levels
Hybrid	Charter planes	Pay for duration, short to medium term/trip	Some control of infrastructure and depend on provider for service levels for duration
Public Cloud	Commercial airlines	Pay as you use, very short term/trip/user	Almost no control of infrastructure or service levels



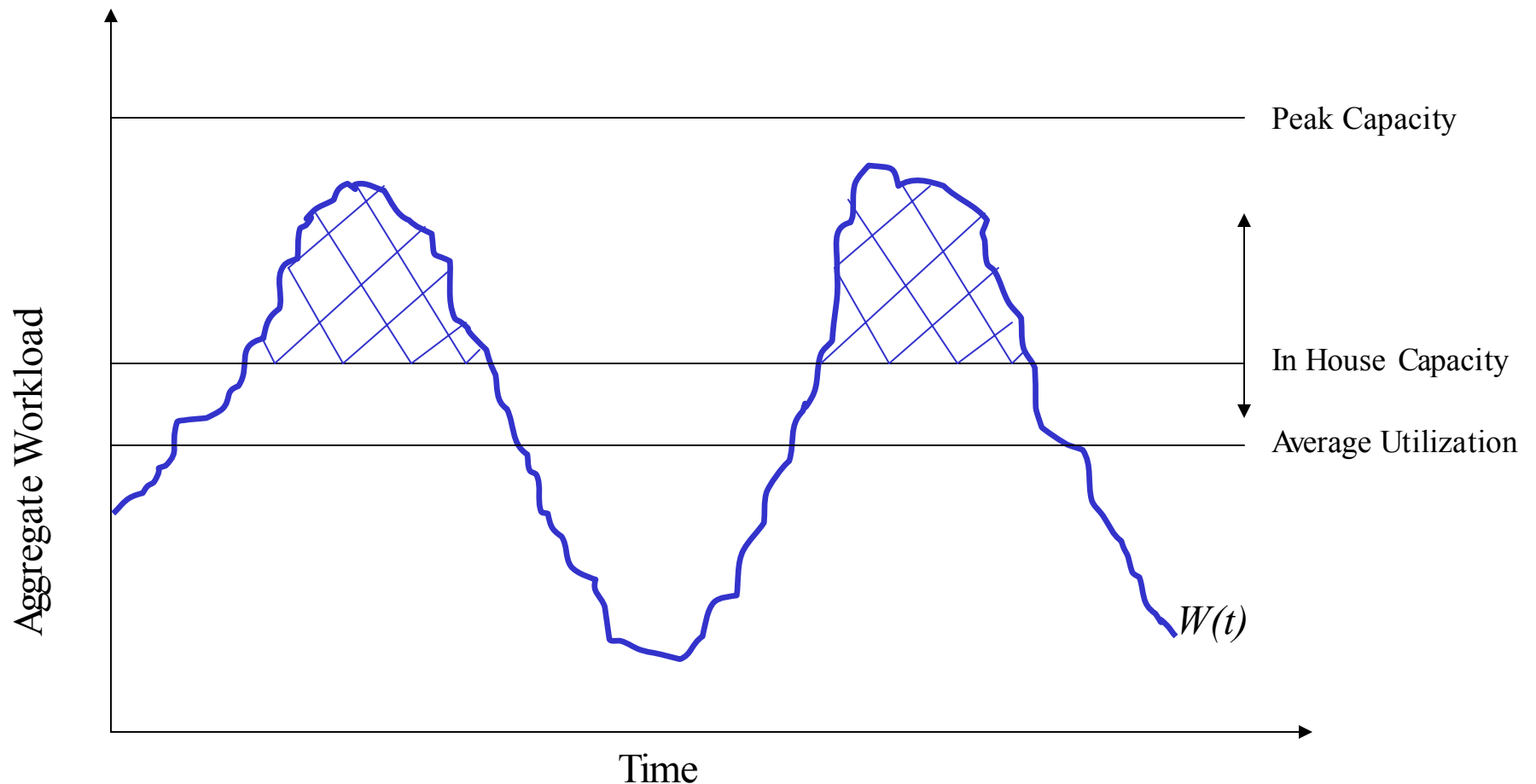
High level IT workload classification by application



Bubble dimension indicative of aggregate IT workload size measured in MIPS or FLOPS

Trends indicated by the arrow are qualitative based on several recent public announcements on cloud computing and other industry research on workload characteristics

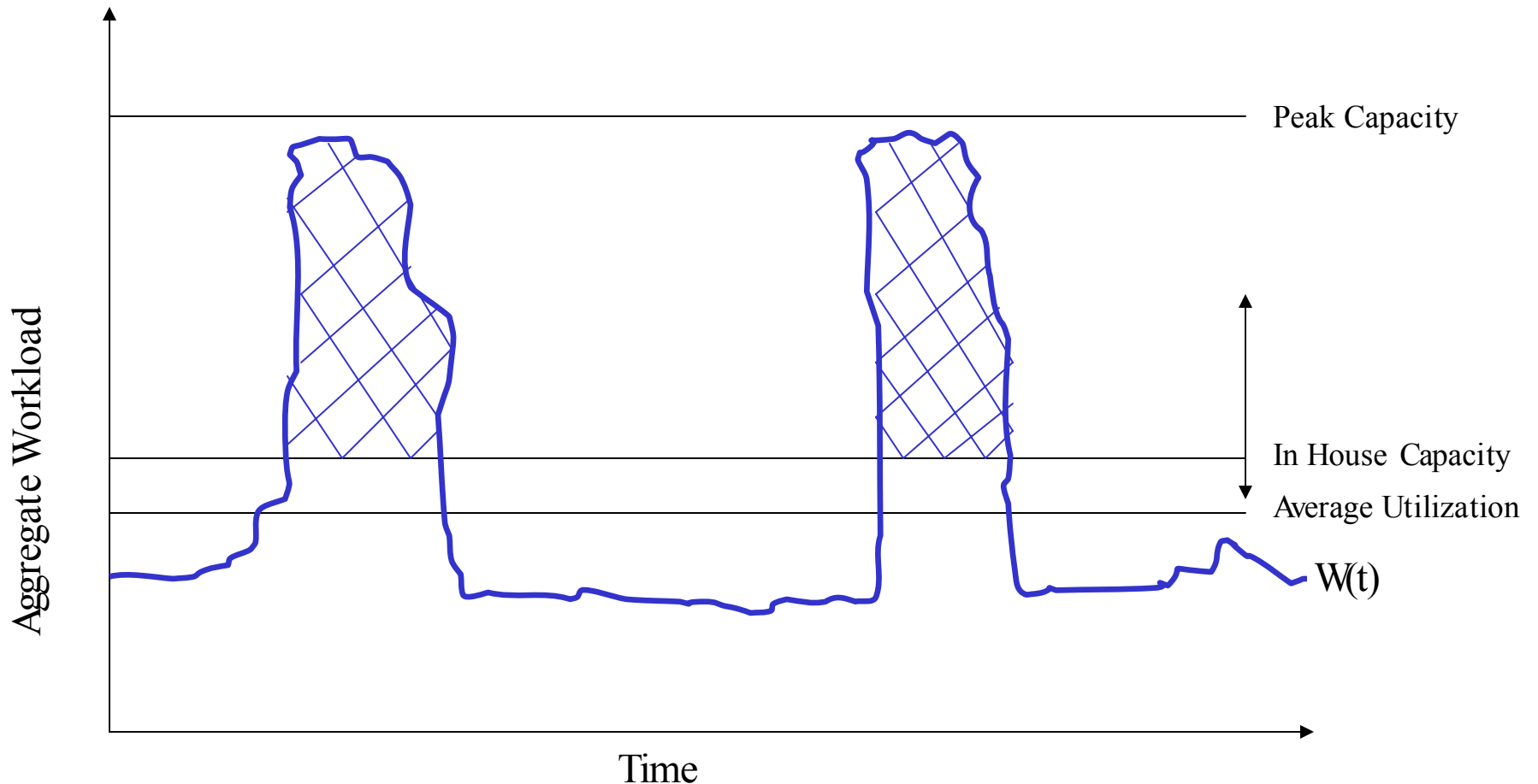
Typical low variability workload graph – common in enterprise business applications



- Shaded area represents under provisioning of IT resources that could lead to dissatisfied users and adverse business impacts
- In house capacity designed for peak results in some over provisioning and wasted resources

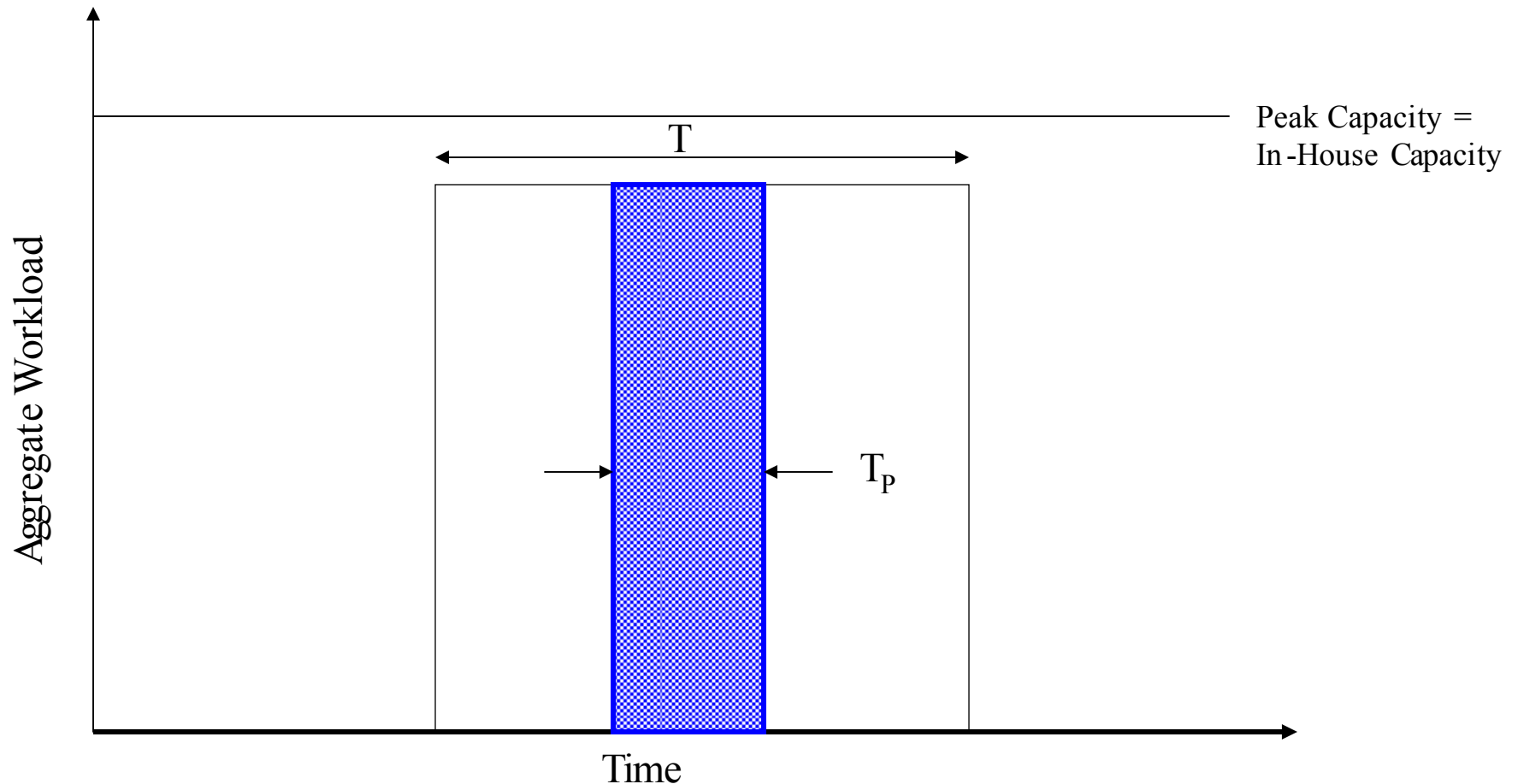
Pulsed (high variability) workload graph – common in departmental or SMB HPC

11



- Shaded area represents under provisioning of IT resources that could lead to dissatisfied users and adverse business impacts
- In house capacity designed for peak results in excessive over provisioning and wasted resources

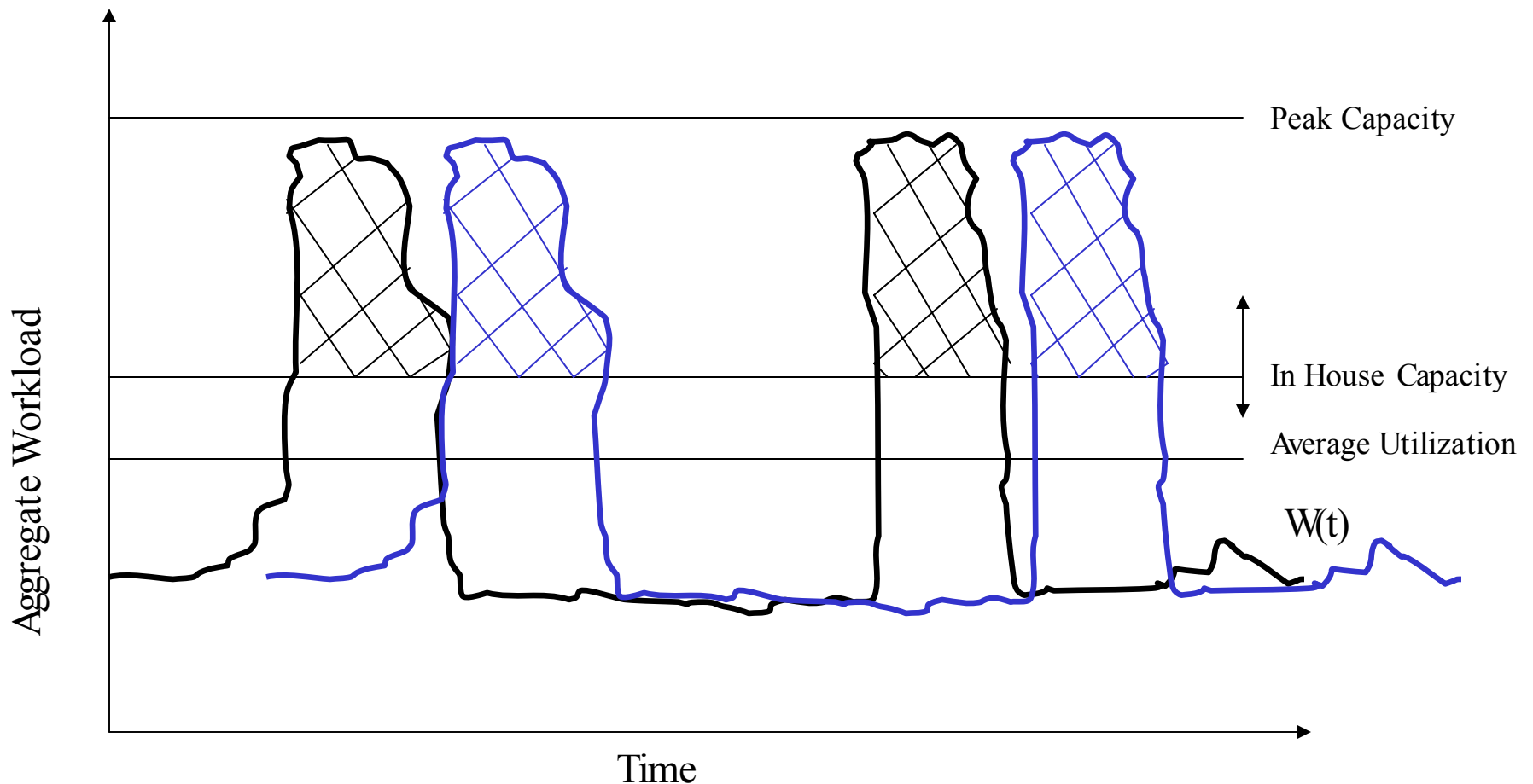
Single Short Pulsed Workload Graph



- Flat Pulsed HPC Workload Occurs When One Large Intensity Workload is Scheduled and Perfectly Balanced
- In house capacity = Peak Capacity and Utilization is Approximately T_p / T , where T_p is the Duration of the Pulse

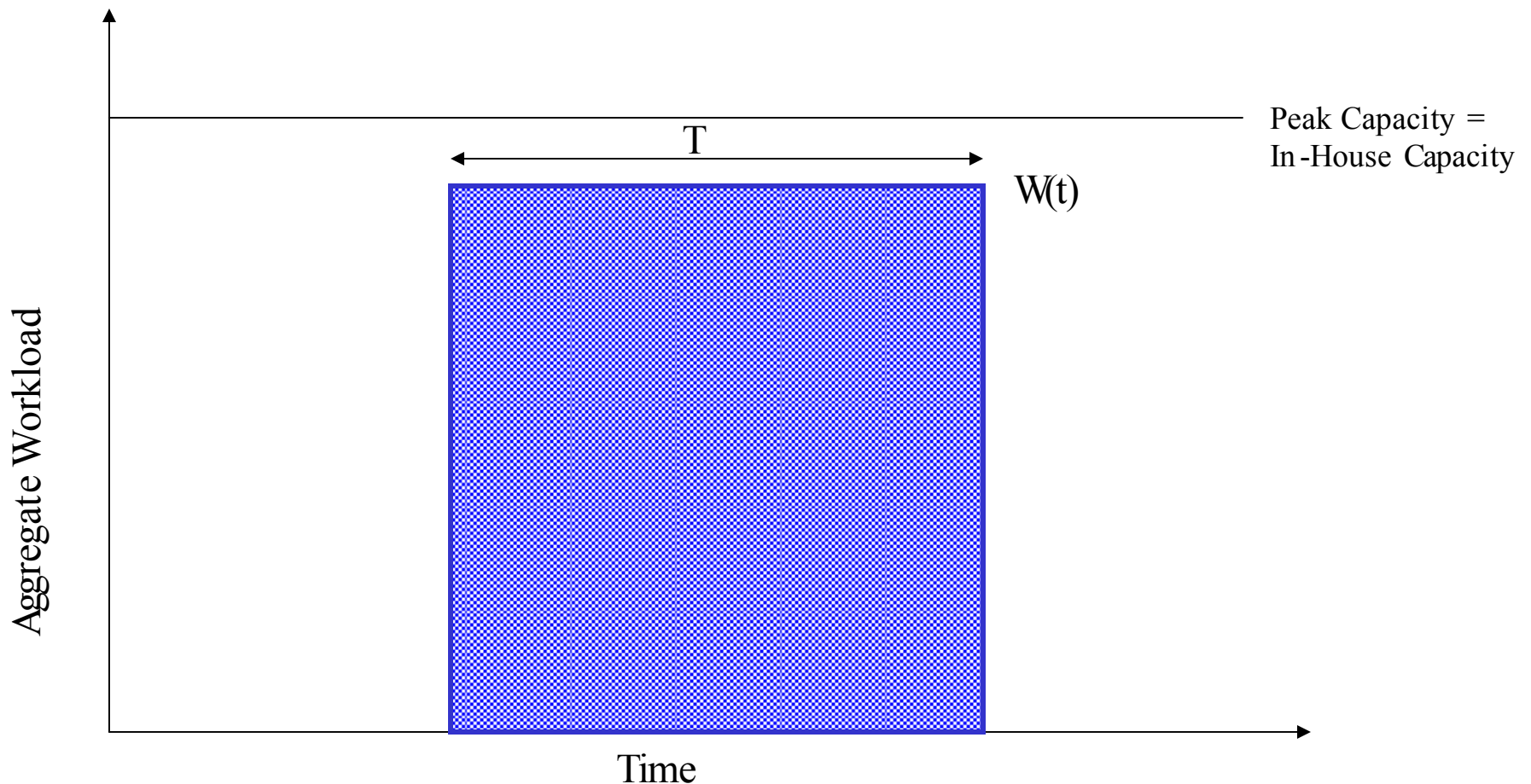
Many pulsed (medium variability) workload graph – common in large enterprise HPC

13



- Shaded area represents under provisioning of IT resources that could lead to dissatisfied users and adverse business impacts
- In house capacity designed for peak results in some excessive over provisioning and wasted resources but average utilization improves compared with the single pulse case

Ideal flat HPC large enterprise workload



- Ideal Flat HPC Workload Occurs When Many Equal Intensity Pulses Coalesce with Perfect Sequencing and Scheduling of the Workload. Illustrates an Ideal Large Enterprise HPC Case.
- In house capacity = Peak Capacity and Utilization is Maximized

Cost – Benefit framework for HPC and Cloud Computing **15**

Client interviews across several industries, industry and workload analyses are needed to examine the benefits.

□ Benefits

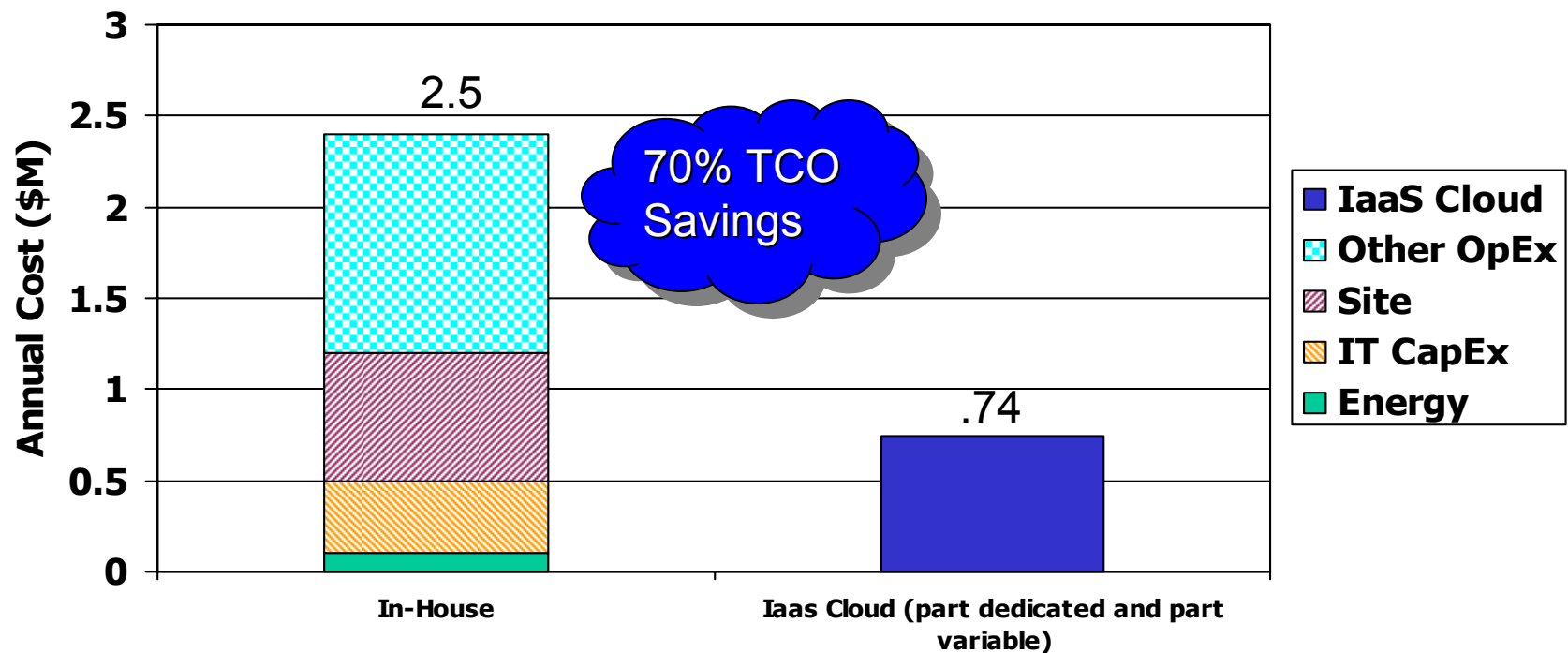
- Business Value: e.g. customer revenues, new business models, compliance regulations, better products, increased business insight, and new breakthrough capability,
- Operational Value: e.g. faster time to results, more accurate analyses, more users supported, improved user productivity, better capacity planning,
- IT Value: e.g. improved system utilization, manageability, administration, and provisioning, scalability, reduced downtime, access to robust proven technology and expertise.

□ Costs

- Data Center Capital Purchases Avoided e.g. new servers, storage, networks, power distribution units, chillers, etc.
- Data Center Facilities Not Built e.g. land, buildings, containers, etc.
- Operational Costs: e.g. labor, energy, maintenance, software license, etc.
- Other Costs: e.g. deployment and training, downtime, bandwidth, etc.

TCO Analysis* for a petroleum services company

Dual Core x86 Cluster System with 384 CPUs (sockets) and 1184 GB of Storage.



Assumptions:

1. 224 CPUs dedicated for 52 weeks at \$.32/CPU hour
2. 140 CPUs variable for 12 weeks at \$.49/CPU hour
3. Includes annual storage and fixed annual contract cost
4. Minimal personnel assumed for 3 shifts for In-House system

* Jonathan Koomey, "A Simple Model for Determining True Total Cost of Ownership for Data Centers", White Paper, The Uptime Institute, 2007.

* TCO model adapted by Cabot Partners

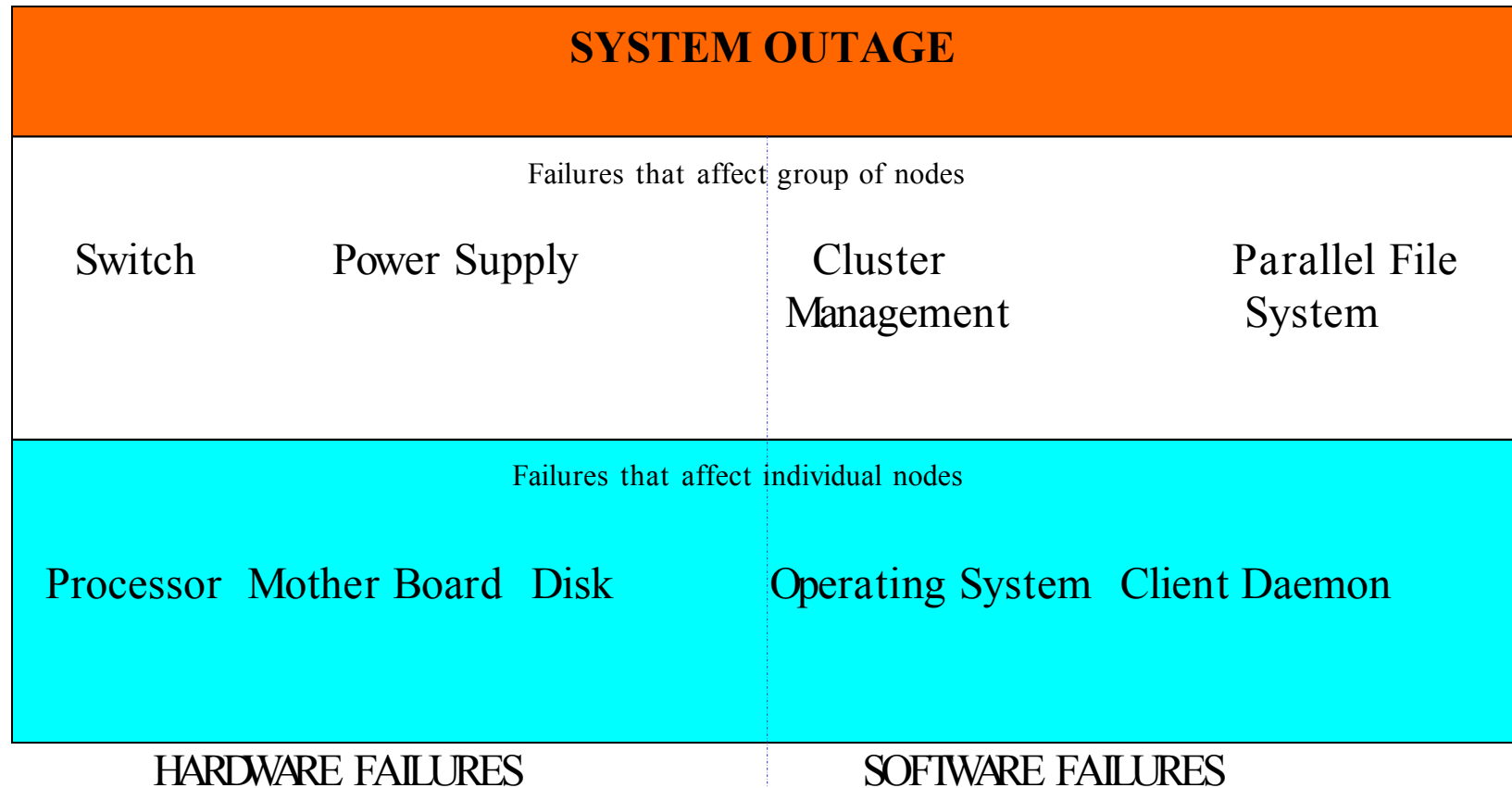
*Availability of Current HPC Systems**



- ❑ Today's supercomputers typically need to reboot to recover from a single failure
- ❑ Entire systems go down (regularly and unscheduled) for any maintenance or repair (MTBF = 40-50h)
- ❑ Compute nodes sit idle while their head node or one of their service nodes is down
- ❑ Availability will get worse in the future as the MTBF decreases with growing system size
- ❑ *Why do we accept such significant system outages due to failures, maintenance or repair?*

*Stephen L. Scott and Christian Engelmann, "Advancing Reliability, Availability, and Serviceability for High Performance Computing", Oak Ridge National Laboratory, April, 2006.

Description of failures of HPC systems*



*Raju Gottumukkala, et. al., "Reliability Analysis of HPC Clusters",
http://xcr.cenit.latech.edu/hapcw2006/program/papers/hapcw_rel_analysis.pdf

Availability Measured by 9's*

9's	Availability	Downtime/Year	Examples
1	90.0%	36 days, 12 hours	Personal Computers
2	99.0%	87 hours, 36 min	Entry Level Business
3	99.9%	8 hours, 45.6 min	ISPs, Mainstream Business
4	99.99%	52 min, 33.6 sec	Data Centers
5	99.999%	5 min, 15.4 sec	Banking, Medical
6	99.9999%	31.5 seconds	Military Defense

Enterprise-class hardware + Stable Linux kernel = 5+

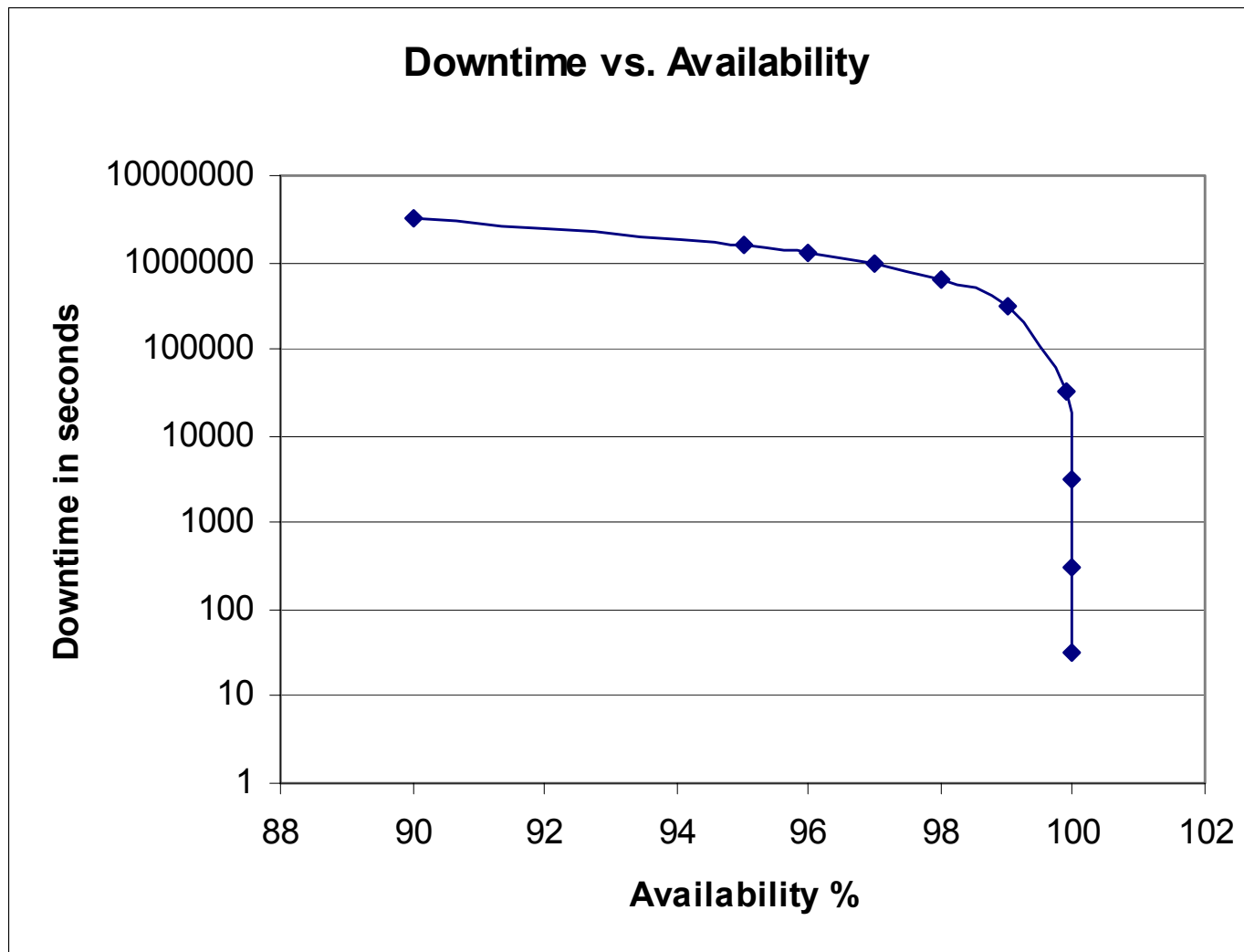
Less expensive hardware + Good high availability package = 2-3

Today's supercomputers = 1+- 3

Desktop = 1 - 2

*Stephen L. Scott and Christian Engelmann, "Advancing Reliability, Availability, and Serviceability for High Performance Computing", Oak Ridge National Laboratory, April, 2006.

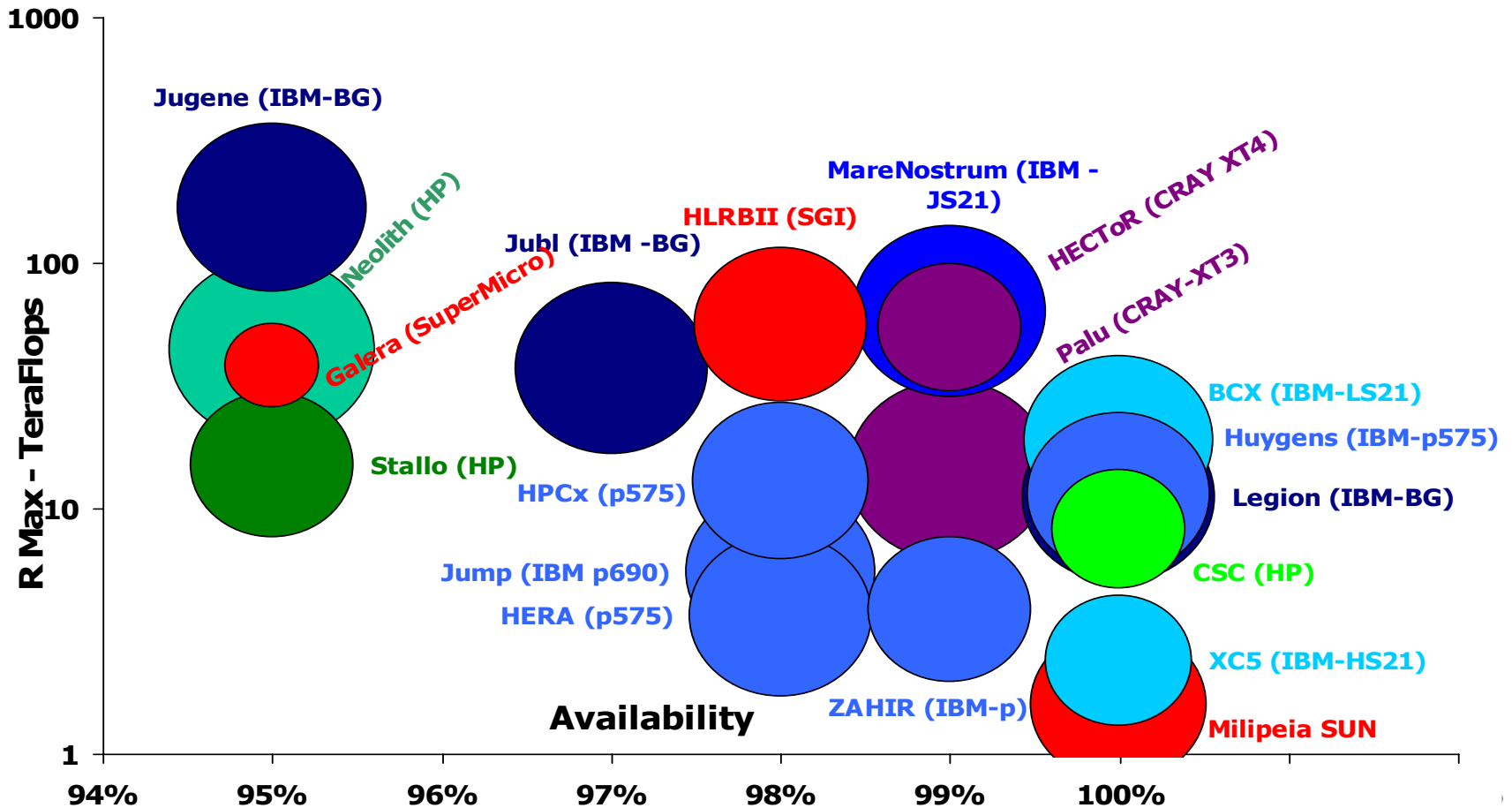
The Downtime reduces exponentially with an incremental improvement in Availability



Failure analysis obtained from system log files and surveys over a long period of time

- ❑ Software failures could be more common as hardware gets more reliable
- ❑ MTBF – Mean Time Between Failures
- ❑ $MTBF_{\text{component}}$ for Hardware Components from manufacturer
- ❑ $A_{\text{component}}$ -- Availability of component = $MTBF_{\text{component}} / (MTBF_{\text{component}} + MTTR)$
- ❑ $MTBF_{\text{system}}$ for entire system aggregates hardware and software
- ❑ Mean Time To Recovery (MTTR)
- ❑ System Availability = $MTBF_{\text{system}} / (MTBF_{\text{system}} + MTTR) \sim 1 - (1 - A_{\text{component}})^n$
- ❑ 4 9's implies 99.99 availability – enterprise class mission critical reliability for data centers
- ❑ Log data from major HPC European installations indicate HPC systems range from 95% to 100% availability
- ❑ For HPC one has to look at Availability/TFlops – larger systems have lower availability as failures are more probable. So we plot performance w.r.t availability and also examine the utilization


Linpack Performance (log values) vs. Availability for Various European Supercomputers - Bubble Sizes Indicative of System Utilization (larger is greater utilization)*



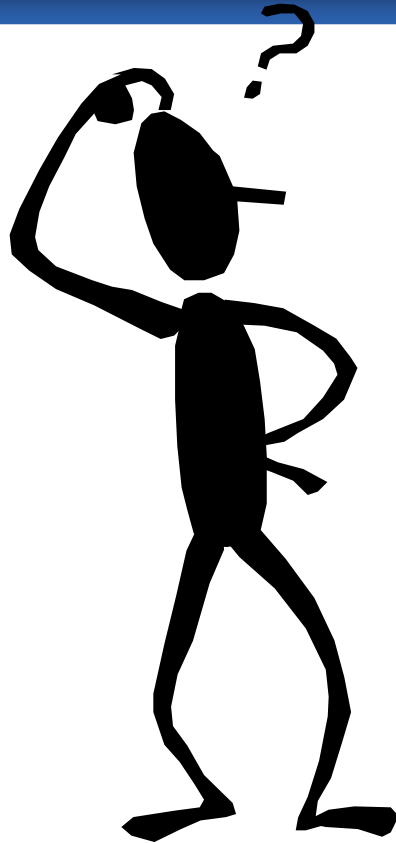
*Alan Simpson, Mark Bull, and Jon Hill, "Identification and Categorization of Applications and Initial Benchmarks Suite", PRACE Consortium Partners, 2008.

Conclusions – Several Related Factors Impact True Cost of Ownership in HPC Data Centers

23

- 
- Performance
 - Capital Acquisition Cost
 - Software Licensing
 - Utilization
 - Availability
 - People Costs
 - Energy
 - Facilities

Questions?



Copyright © 2009. Cabot Partners Group, Inc. All rights reserved. Other companies' product names or trademarks or service marks are used herein for identification only and belong to their respective owner. The information and product recommendations made by the Cabot Partners Group are based upon public information and sources and may also include personal opinions both of the Cabot Partners Group and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. The Cabot Partners Group, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.