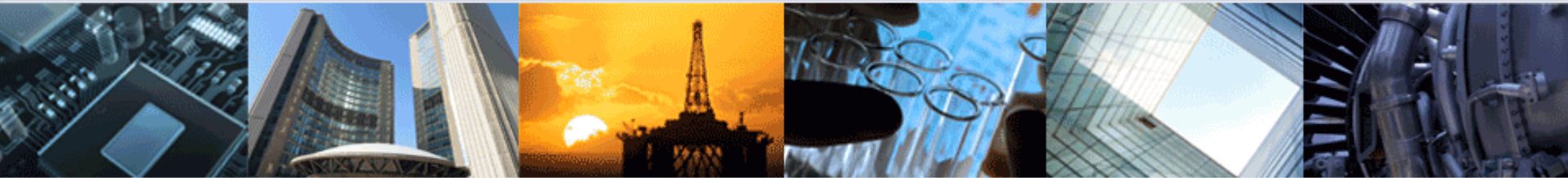# Virtualization & Grid

Presented by: Bernhard Schott, bschott@platform.com
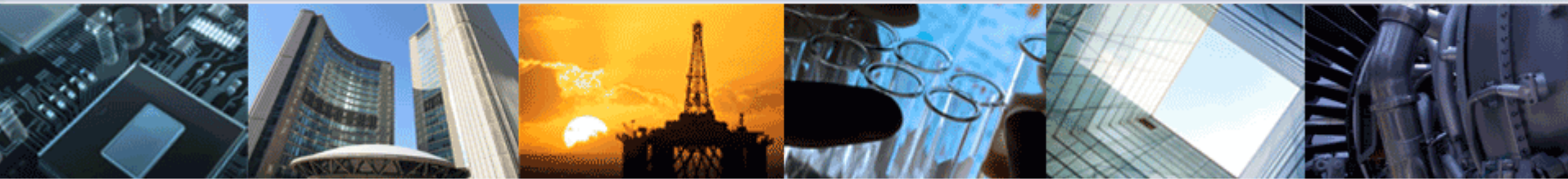
Date: 27th May 2008

Agenda

- Platform Computing

- Team play: Grid & Virtualization

- Business process speedup

  – Example: "self services" on VMs and complex infrastructures

- Virtualization Technologies & Advanced Use Cases

- Summary

Platform™ Introduction

Platform Computing

Platform is a pioneer and the global leader in High Performance Computing infrastructure software, delivering integrated software solutions that enable organizations to improve time-to-results and reduce computing costs.

Over 2200 Customers Worldwide

- Electronics, Financial Services, Manufacturing, Life Sciences, Oil & Gas, Government, Universities & Research, Telco …

Recognized Leader in HPC, Cluster and Grid Computing

- 15 years global experience
- Worldwide offices, resellers and partners
- 24x7 follow the sun support and services

Growing & Profitable since inception in 1992

- Self-funded
- No debt; money in bank

Platform

Office Locations

North America
Toronto (HQ)
San Jose
Washington
Detroit
Los Angeles
Boston
New York

International
China
Japan
Korea
UK
Germany
France

VARs

U.S.	South Africa
Italy	Sweden
Israel	Turkey
Germany	India
Spain	Malaysia
Korea	Thailand
Taiwan	Australia
Singapore	Austria
Japan	France
U.K.	China
The Netherlands	U.A.E.
Poland	Portugal
Brunei	

Platform™   Big Companies Trust Us

Financial Services
- BNP Paribas
- Citigroup
- Deutsche Bank
- Fortis
- HSBC
- JP Morgan Chase
- Lehman
- Mizuho Financial
- MUFG
- Prudential
- Société Générale
- Sal Oppenheim

Electronics
- AMD
- ARM
- ATI
- Broadcom
- Cadence
- Cisco
- HP
- IBM
- Motorola
- NVIDIA
- Qualcomm
- Samsung
- ST Micro
- Synopsys
- TI
- Toshiba

Industrial Manufacturing
- Airbus
- BMW
- Boeing
- Bombardier
- British Aerospace
- Daimler & Chrysler
- GM
- Lockheed Martin
- Pratt & Whitney
- Toyota
- Volkswagen
- Xi'an Aircraft Design

Life Sciences
- AstraZeneca
- Bristol Myers-Squibb
- Celera
- Dupont
- GSK
- Johnson & Johnson
- Merck
- Novartis
- Novo-Nordisk
- Pfizer
- Wellcome Trust Sanger Institute
- Wyeth

Government & Research
- ASCI
- CERN
- CINECA
- DoD, US
- DoE, US
- ENEA
- ETH
- Fleet Numeric
- GSI
- INFN
- MaxPlanck
- SSC, China
- TACC
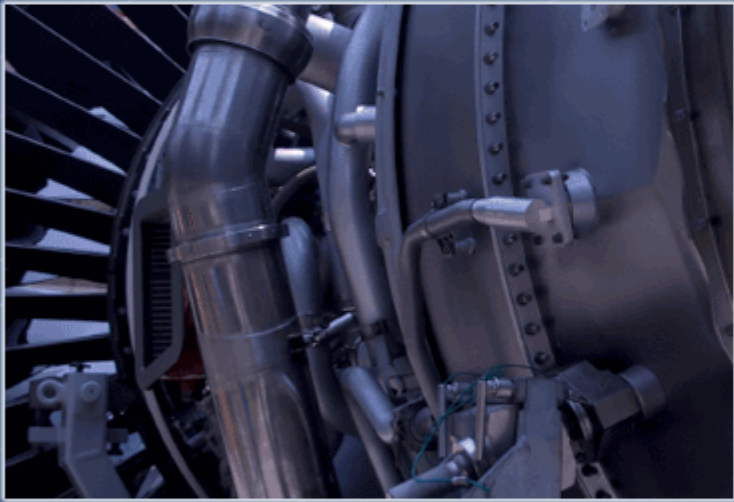- TU Dresden
- Univ Tokyo

Other Business
- Bell Canada
- Cablevision
- Deutsche Telekom
- Ebay
- Starwood Hotels
- Telecom Italia
- Telefonica
- Sprint
- GE
- IRI
- Cadbury Schweppes

Partner Ecosystem

Strategic Partners

Premier Partners

Select Partners

Business Benefits

Platform

Developer   Executive   User   ISV   Executive   I.T. Manager

Platform Enterprise Solution

Platform Accelerate

Make money
Grow revenue

1. **Do more, faster**
2. **Reduce time to market**
3. **Simplify use of applications**
4. **Scalable, grid-enabled apps**
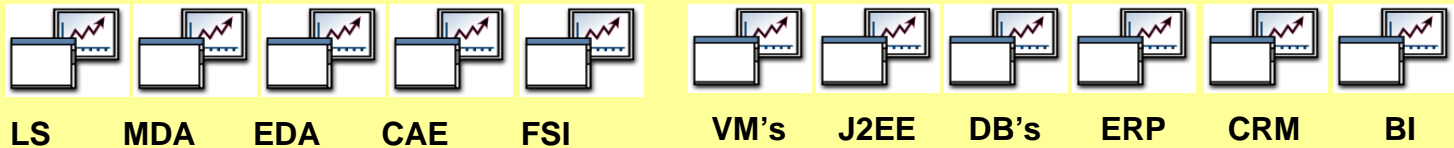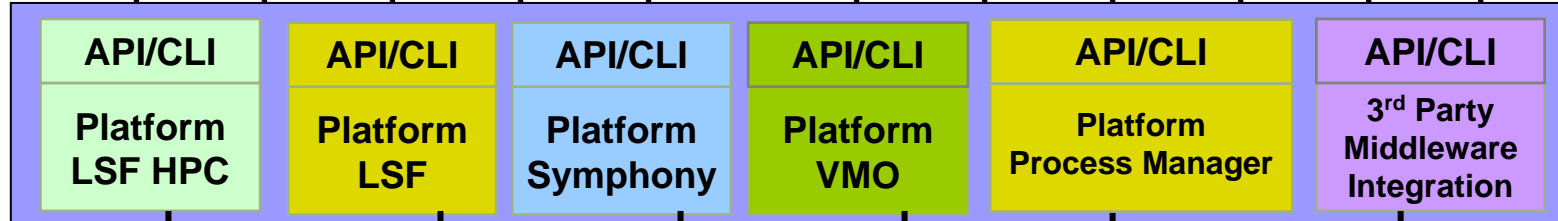5. **Improve staff productivity**

Platform Manage

Save money
Reduce costs

1. **Greater control of infrastructure**
2. **Reduce TCO**
3. **Save power & cooling**
4. **Simplify HPC management**

Scalability Roadmap
Based on projected Customer Cluster Growth

# of Sockets

Based on Dual Socket Hosts

"Spruce"

"Cypress"

"Hawthorn"

LSF 7 "Ironwood"

The grey shaded area is our projected design cone.

The efficiencies experienced by our customers with multi-core processors will define where we will be in this cone.

- **LSF, built on EGO, will increase the scalability of LSF from thousands of nodes to tens of thousands of nodes**

Platform™
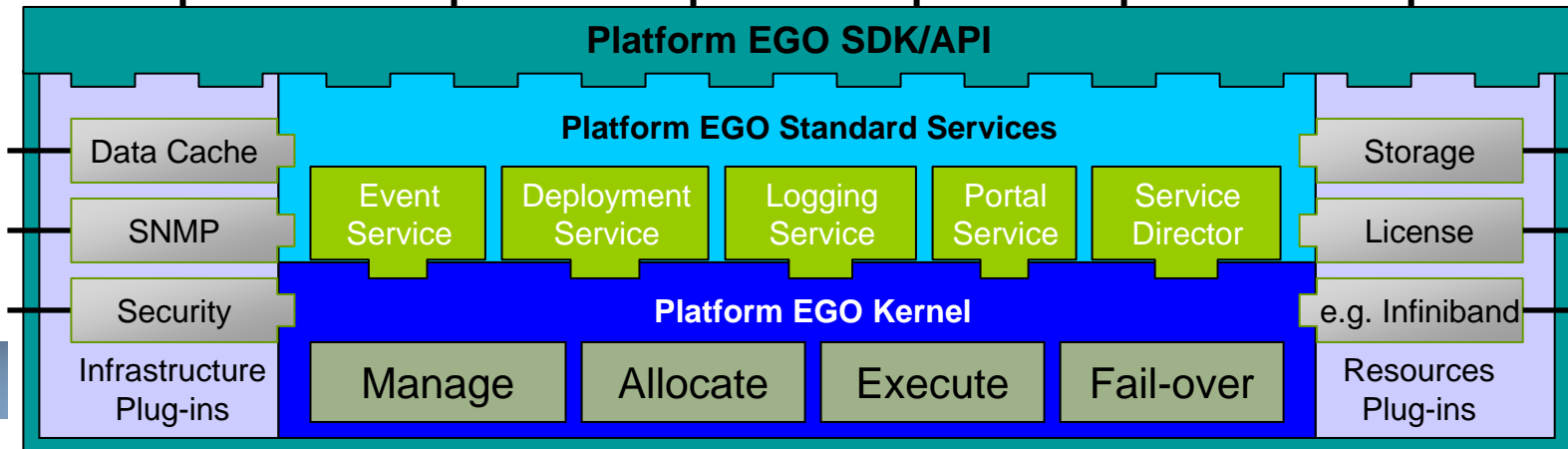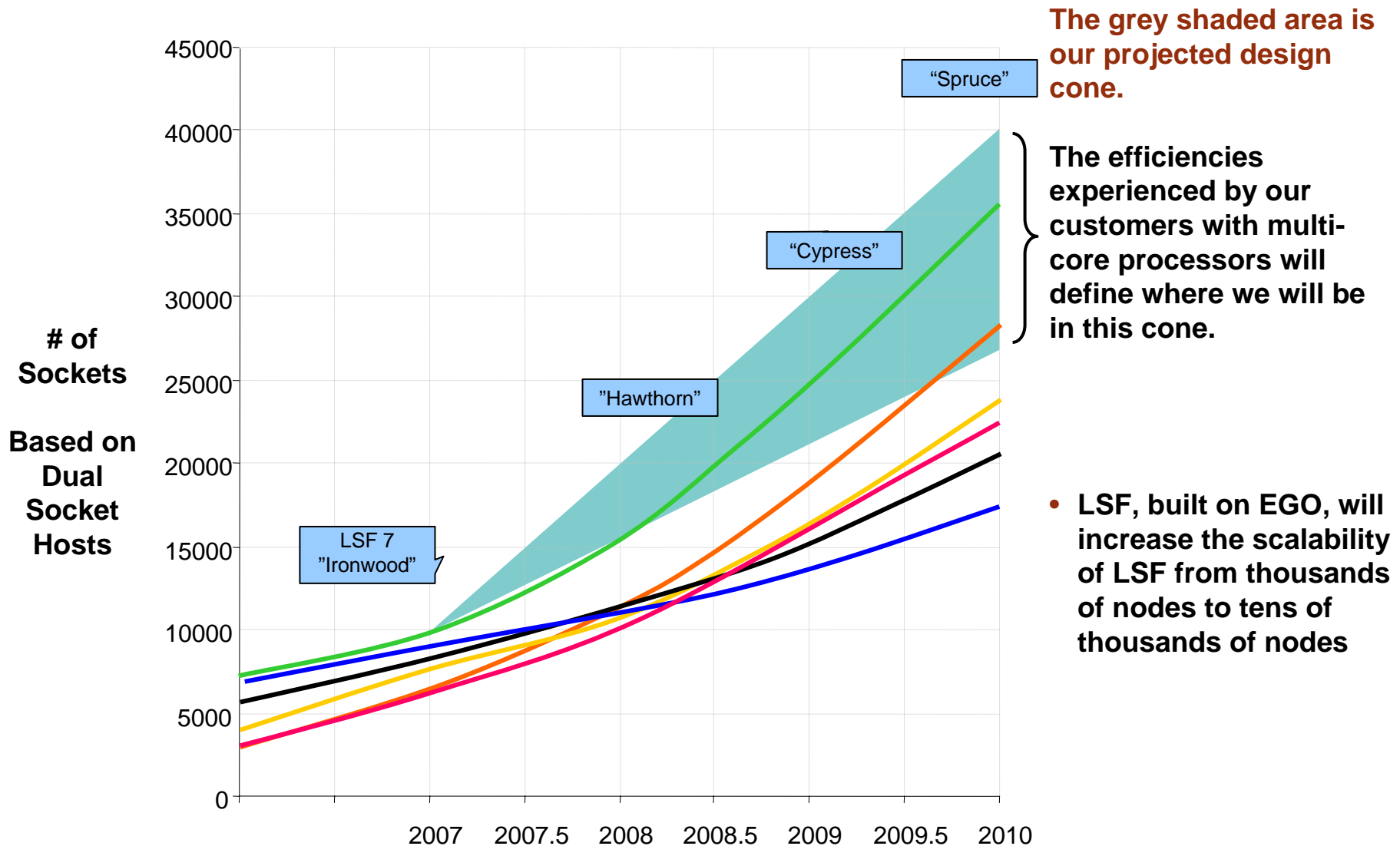
- ## Performance:
  - 10millions jobs per day throughput with >95% job-slot utilization based on EDA job mix, max 5min for failover. (EDA job mix: 5min, 15min, 30min job-runtime)

- ## Scalability (LSF7.0):
  - n*1000's users, 7500 hosts (up to 10000 hosts dep. on load profile),
  - >100Hz sustained submission rate, (much higher with LSF7.0.3)
  - 0,5 million jobs in single logical LSF cluster at any time
  - 8196's way-parallel jobs with full control down to each single task

- ## Reliability: self-healing, recovery from incidents, policy driven proactive problem containment, **no** job loss during operation or in error condition, reconfiguration or failover

- ## Compatibility: Implemented OGF Standards
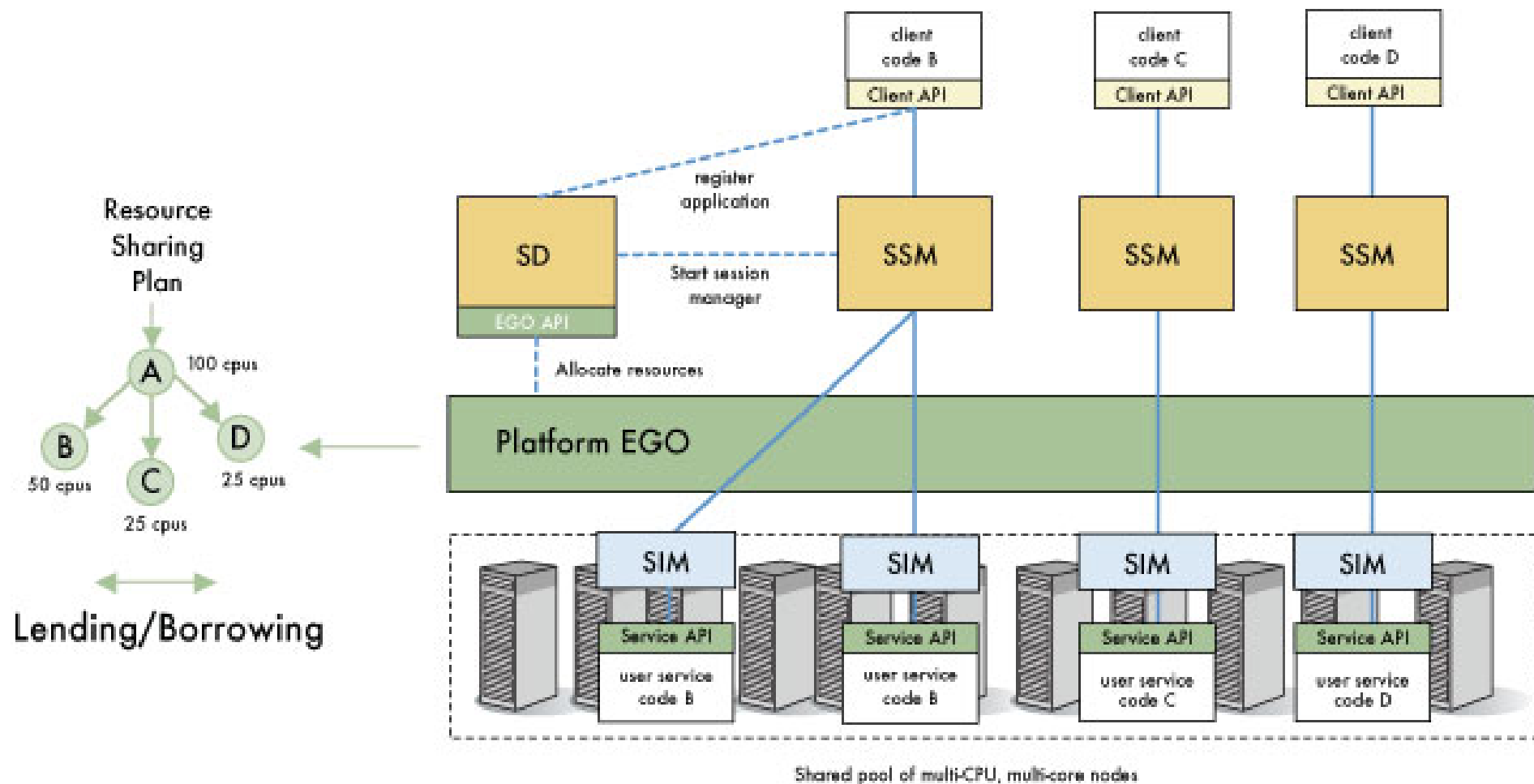 - JSDL / HPC-Basic Profile +extensions / BES
  - DRMAA

Quality requirements

- Performance and Scalability translates into Reliability

- Reliability can be measured as "MTBF" - Mean Transactions (=Jobs) Between Failure

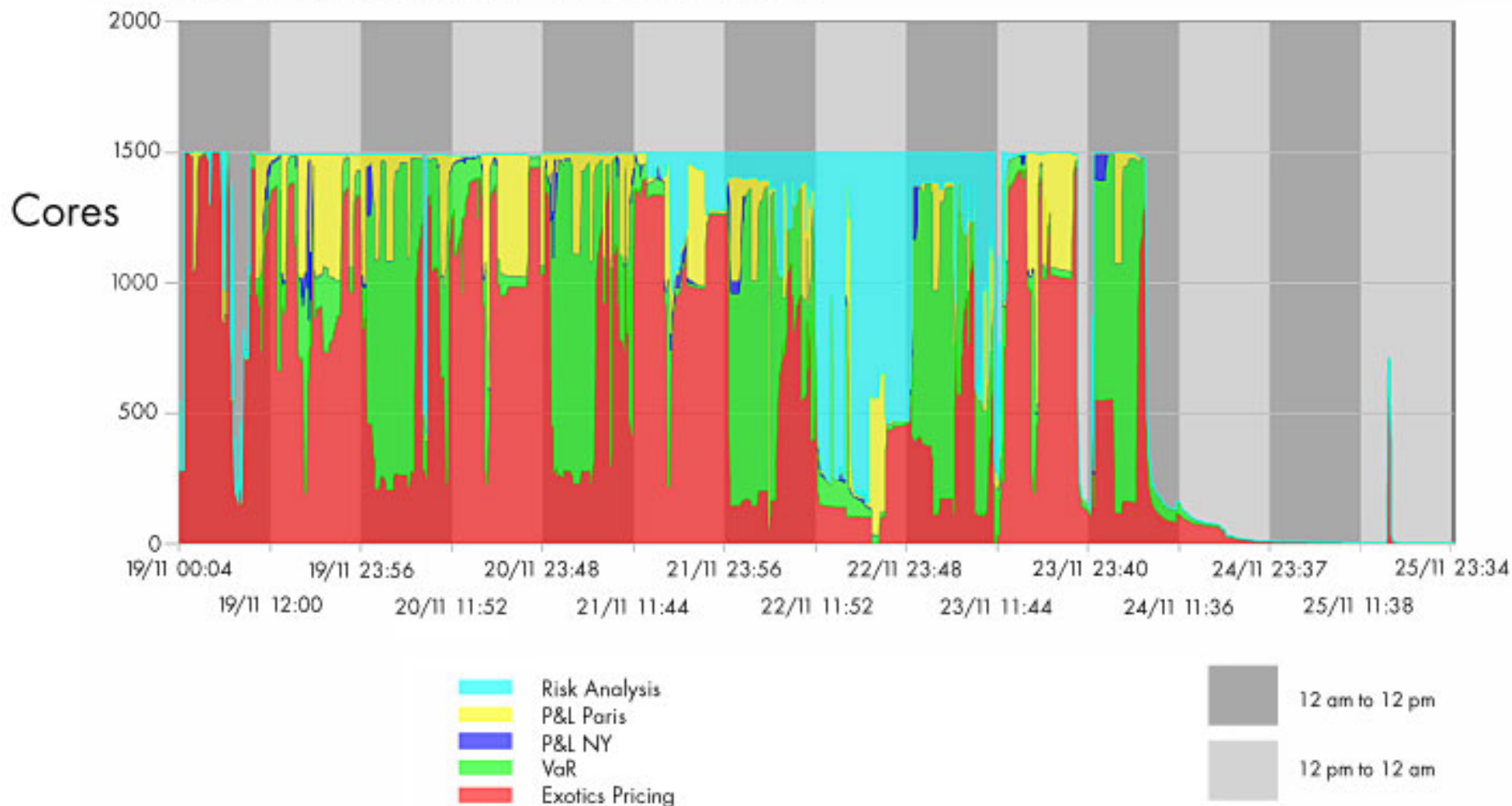- Reliability is the prime requirement driving Platform technology

© Platform
Computing Inc.

Performance, at any Scale

Platform

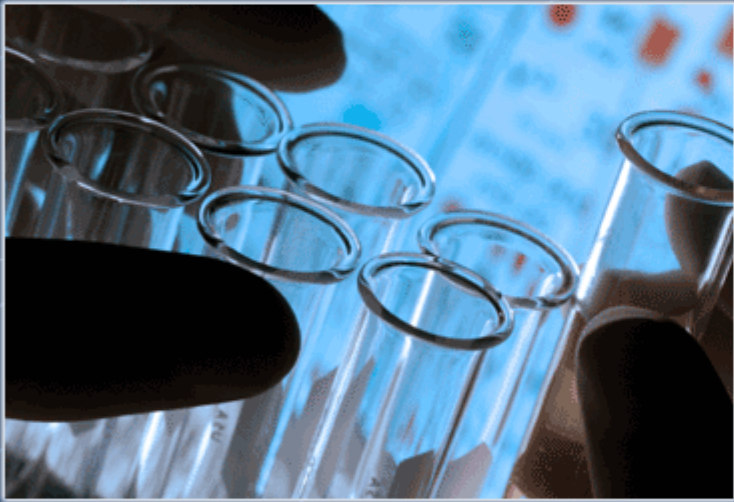|	Symphony at IBM DCCoD
Scalability 1,000 concurrent clients, 100 applications	**20,000+ CPU's** simulated on 1,000 physical CPUs in one cluster
CPU Utilization 1-100 clients, 1 sec task, 1KB message, 2,000 CPU	**98%**
Task Throughput 1KB Message	**2,700 messages/sec**
Single Task Round Trip	**2.4 ms**
Single Session Round Trip 100KB common data, 10 second 1KB Task	**11.8 ms**

© Platform
Computing Inc.

Cluster partition usage over the last week

Report period: from 11-19 00:00 to 11-26 00:00

Legend:
- Risk Analysis
- P&L Paris
- P&L NY
- VaR
- Exotics Pricing

- 12 am to 12 pm
- 12 pm to 12 am

© Platform
Computing Inc.

Platform Process Manager

• Workflow Automation

Platform VMO

Dynamic Compute Grid with on-demand provisioning

Platform

VMO provisions enough VMs to handle the Windows workload

VMO provisions enough VMs to handle the Linux workload

User 1 submits a job requiring a Linux environment

User 2 submits a job requiring a Windows environment

Windows LSF nodes are stopped when no longer needed

VMO Agent
Hypervisor
CPU 60%   Host 1   2GB RAM Avail

VMO Agent
Hypervisor
CPU 60%   Host 2   2GB RAM Avail

LSF User 1

LSF User 2

LSF Grid Master

VMO Admin User defines resource Policies for user VMs sharing physical server hosts

VMO Agent
Hypervisor
CPU 60%   Host 3   2GB RAM Avail

VMO Agent
Hypervisor
CPU 30%   Host 4   4GB RAM Avail

Provisioning manager provisions the appropriate software stack to the VM on-demand.

Platform VMO

Platform EGO

Admin

Web-based Management Console

Shared Storage

OCS/Kusu/ScaliManager Provisioning Manager

VMO Architecture

Platform

VMO Mgmt Console

Virtual Machine | Virtual Machine

Xen Enterprise

VMOAgent
EGOAgent
Hypervisor

VM Storage

VMO Manager

VMO/EGO Persistence

EGO Master

VMO Mgmt Server
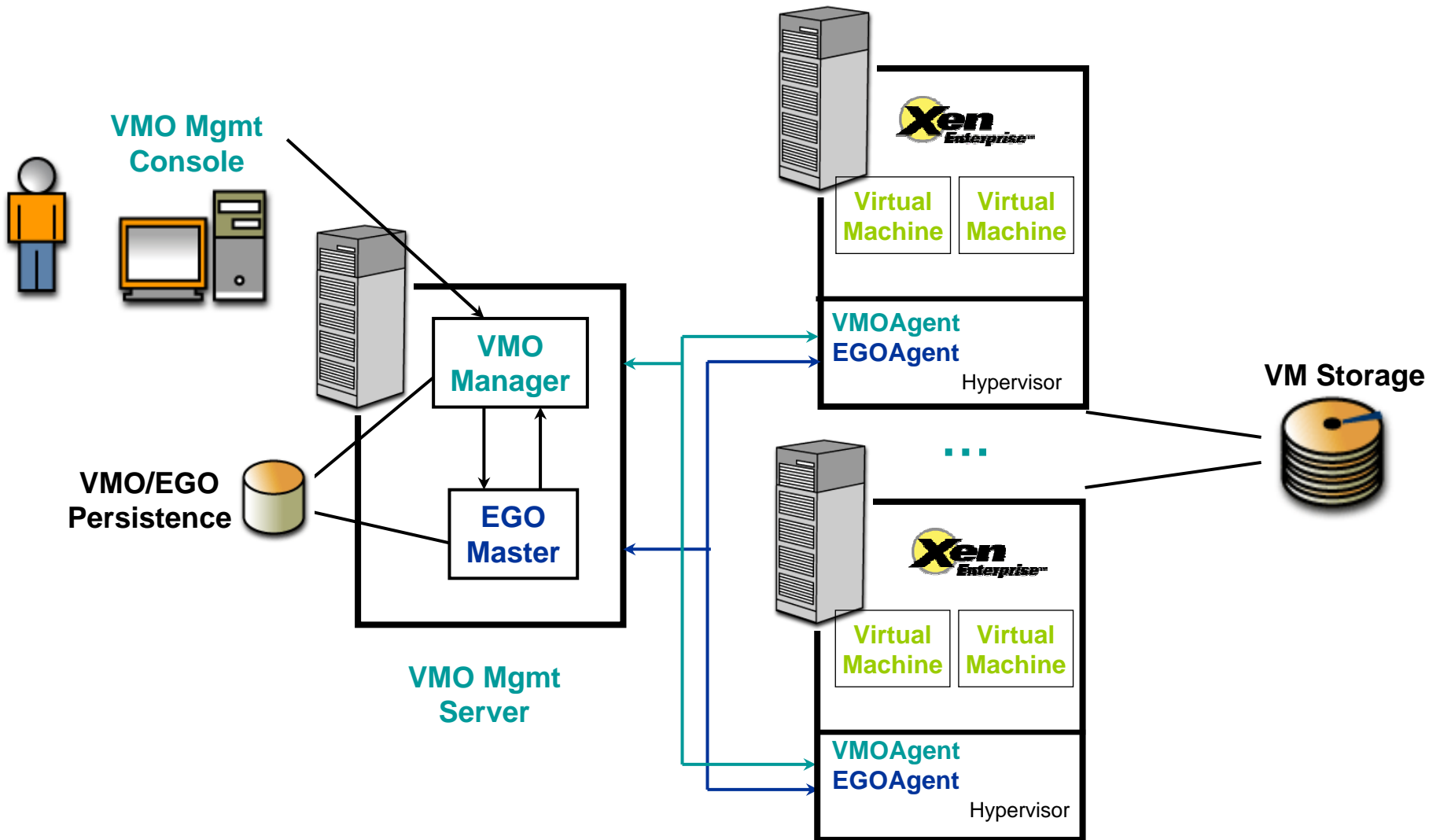
Virtual Machine | Virtual Machine

Xen Enterprise

VMOAgent
EGOAgent
Hypervisor

Business process speedup

- Example use case: "versioning"
- Users: chip designers, CAD engineers
- Challenge: customer driven projects
  - Need to use correct application configuration and version of libraries for each project
  - Need for correct application version and patch level – not the latest, but those defined by the customer!
- IT challenge: this morning, give me
  - 3 hosts RHEL4+Patch#47+App-version#7 + component-lib-version#68
  - Afternoon: … another project … but keep my morning machines until customer calls back

- Example use case: "versioning"
- Solution: "self service"
- Users request VMs with project specific properties by GUI or command line (CLI)
- CLI allows script integration → automation of project specific setups
  - Instead of scrolling through GUI-Listings for correct choices, have an icon on your desktop – 1 click – here you are!
- Tremendous reduction of operational costs
- Better reproducibility – less errors
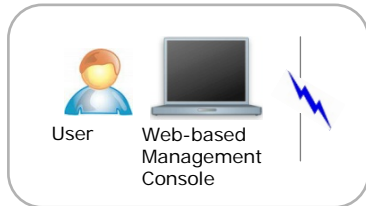
User logs into VM Management Console

User sees only VMs and VM templates that they have been assigned

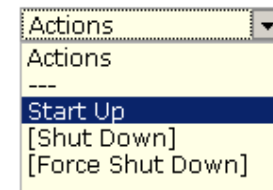User starts up VM with VMO selecting the appropriate host
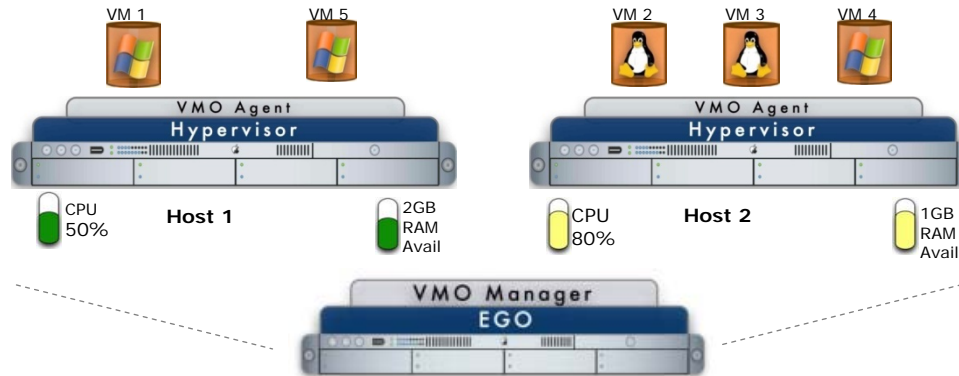
Templates & VMs are stored on shared storage

Admin creates VMO users

Admin creates VMs and VM templates
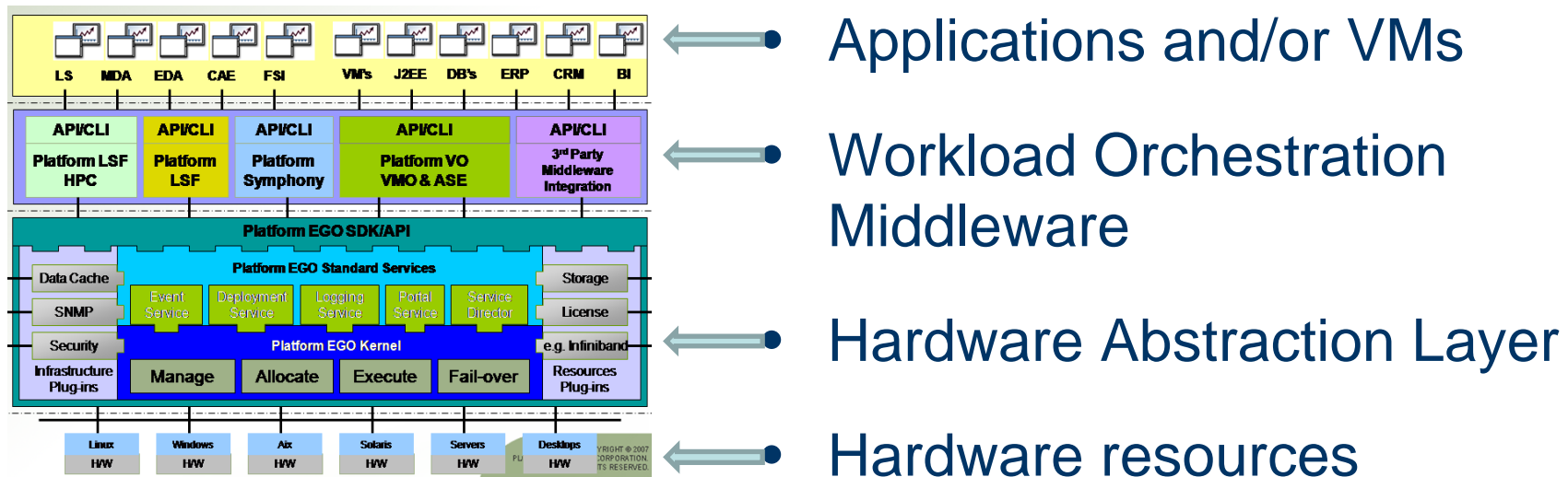
Admin assigns user to VMs and/or VM templates

# Virtualization Technologies & Advanced Use Cases

- ## Virtualization is taking place on several layers using different technologies:
  - ### – Virtual Machines (VM)
  - ### – Grid technology

Applications and/or VMs

Workload Orchestration Middleware

Hardware Abstraction Layer

Hardware resources

- ## Virtual Machines can
  - be launched by the Grid
  - themselves participate in the Grid
  - become dynamic components in the Grid
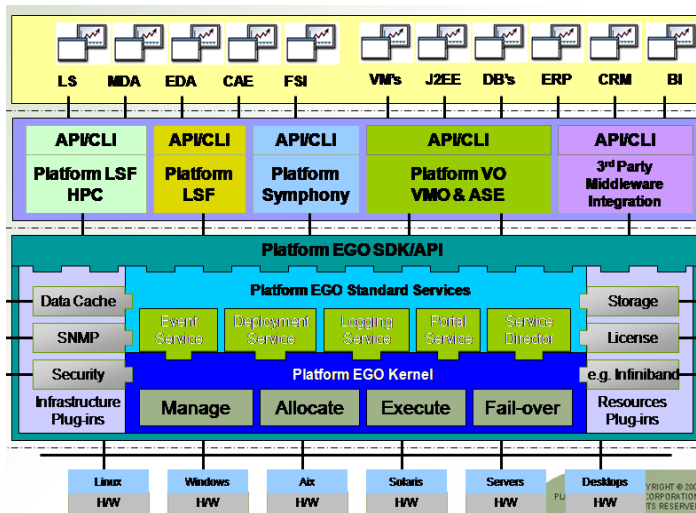
VMs launched by the Grid become resources in the same Grid

- Automated deployment of compute nodes
  - dynamic workload driven integration (growth) and release (shrink) of resources

- Per request provisioning of multi-tier systems

- Multi-Tenant infrastructure:
  - separate Grids per consumer group
  - Fully independent configuration

- On demand multi-site extension
  - Transparent workload transfer to remote sites
  - Integration of resource providers like Amazon, IBM, HP

Platform

LSF 7.0.2 & VMO 3 are configured and running

VMO RRB maps resource strings in LSF to VMs managed by VMO

Job is submitted

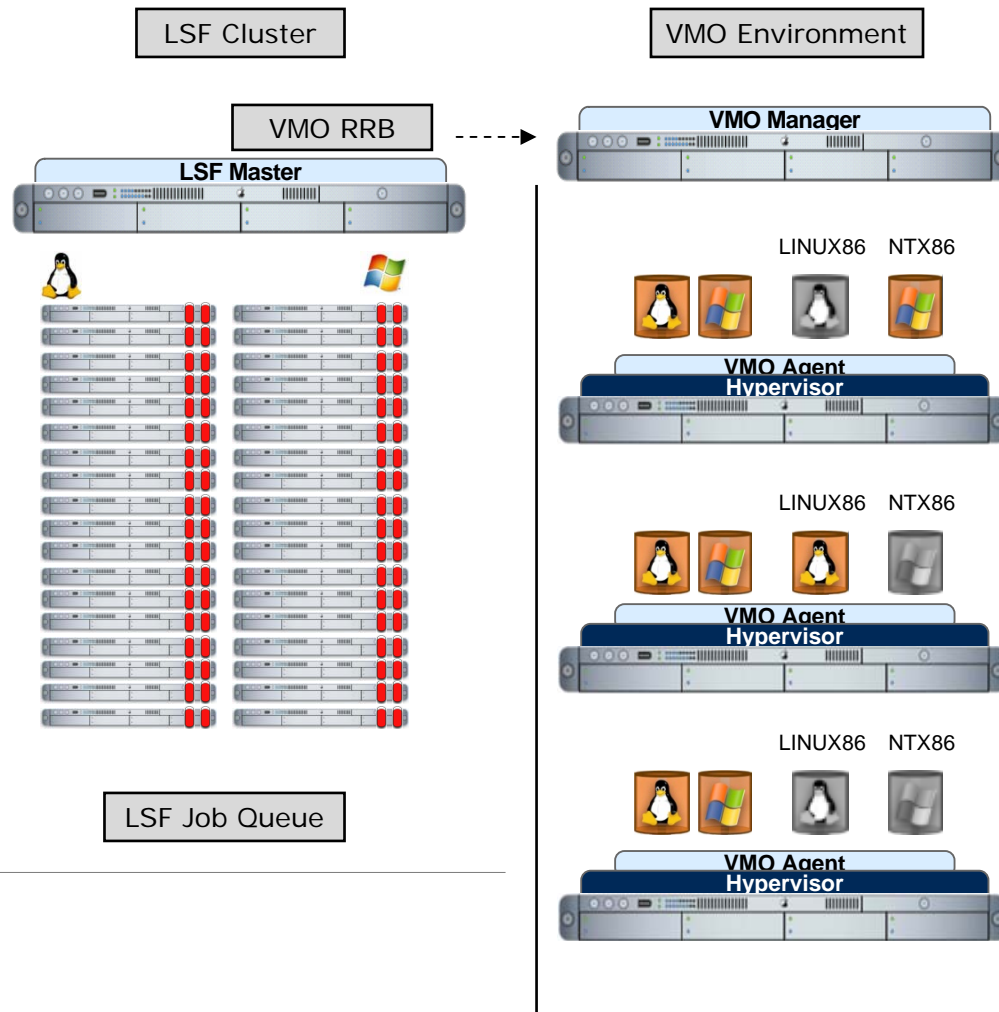If the job persists in the queue for a defined period of time a specific VM is started

There are now resources available, the job is(are) executed
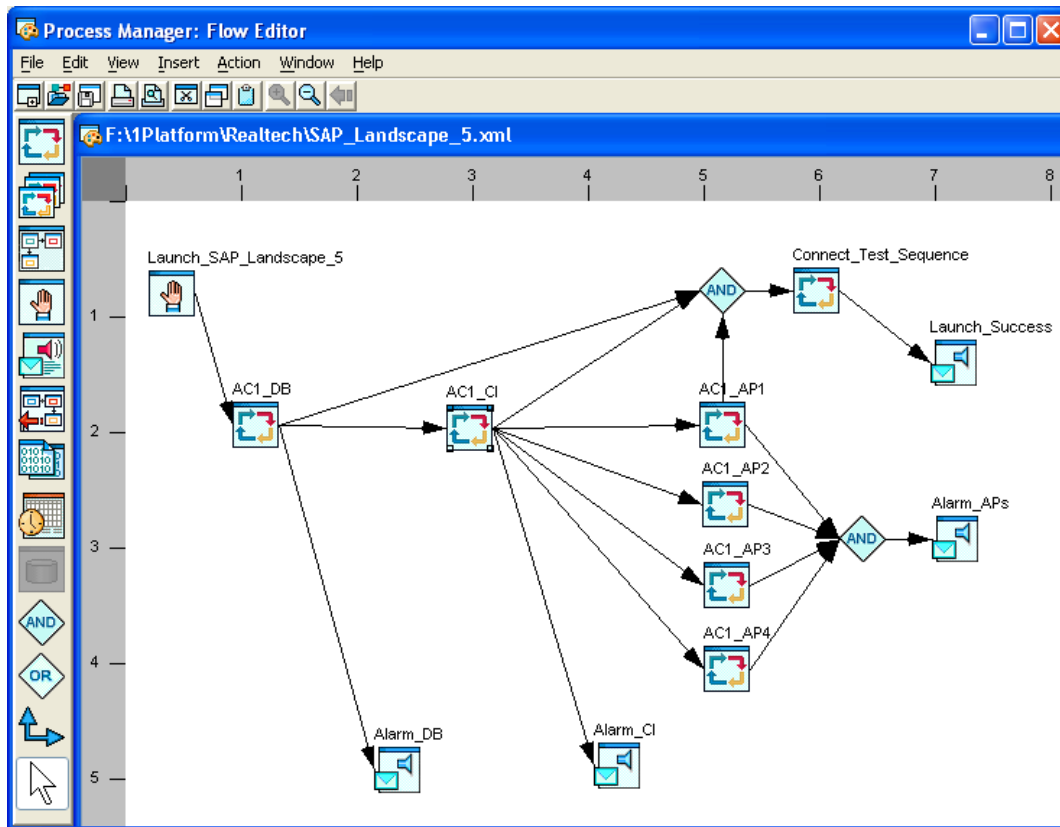
Another job is submitted

If the job persists in the queue for a defined period of time a different, specific VM is started

There are now resources available, the job is(are) executed

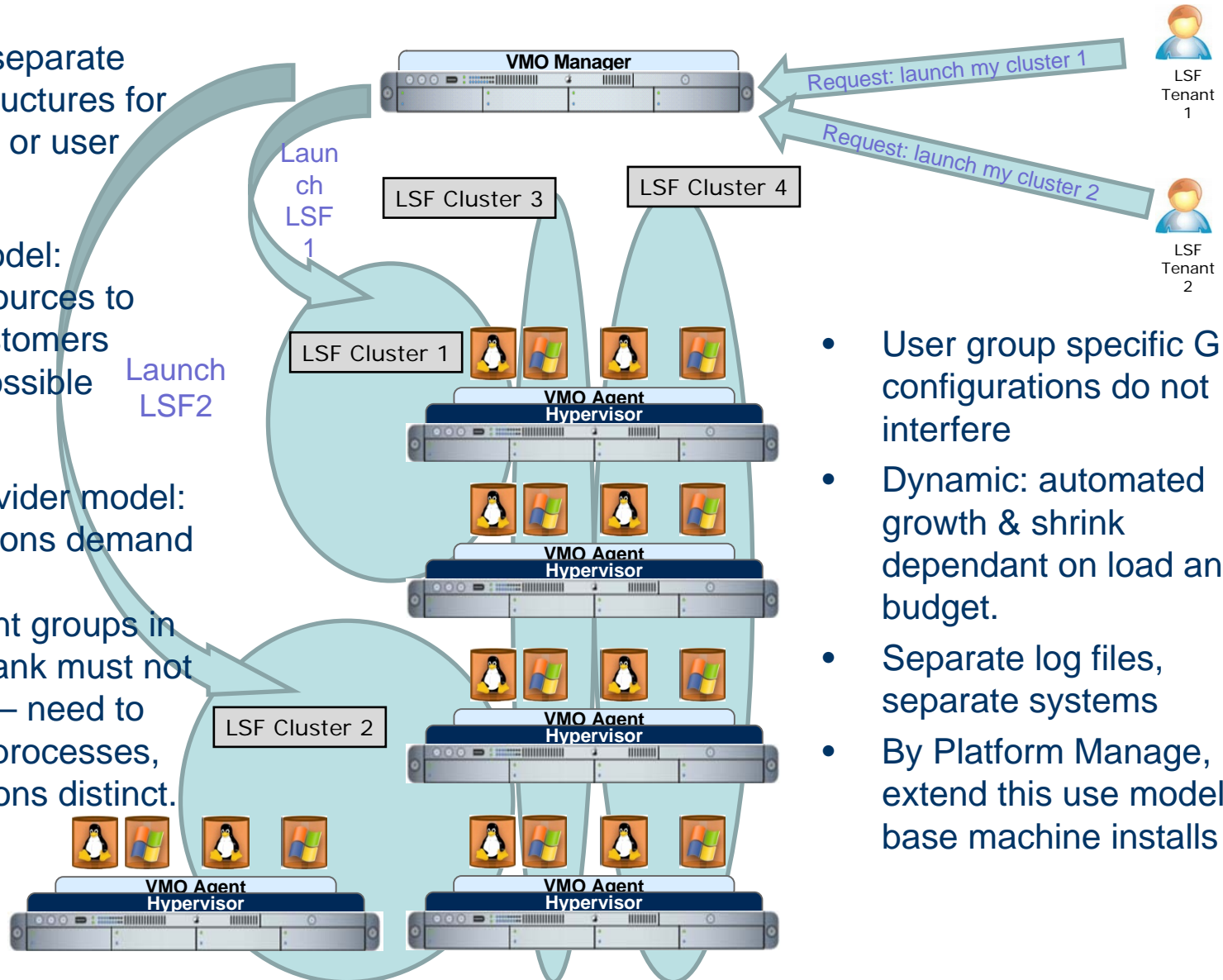When the VMs remain inactive for a defined period of time they are shutdown

LSF Cluster

VMO Environment

VMO RRB

VMO Manager

LSF Master

LINUX86 NTX86

VMO Agent
Hypervisor

LINUX86 NTX86

VMO Agent
Hypervisor

Shared Storage

LINUX86 NTX86

LSF Job Queue

VMO Agent
Hypervisor

- # Provisioning Workflow Automation
 - ## – Multi-tier infrastructure initiation

- Instantiate separate System for Audit
- Temporarily run new (/older) version for compatibility
- Testing of new functionality
- Test upgrade impact
- Training instance / delete after use

- Instantiate separate Grid infrastructures for each tenant or user group

- Provider model: provide resources to multiple customers with best possible separation.

- Internal provider model: e.g. regulations demand certain risk management groups in the same bank must not collaborate – need to keep data, processes, and operations distinct.

VMO Manager

Request: launch my cluster 1

Request: launch my cluster 2

LSF Tenant 1

LSF Tenant 2

Launch LSF 1

Launch LSF2

LSF Cluster 3

LSF Cluster 4

LSF Cluster 1

VMO Agent
Hypervisor

VMO Agent
Hypervisor

VMO Agent
Hypervisor

LSF Cluster 2

VMO Agent
Hypervisor

VMO Agent
Hypervisor

- User group specific Grid configurations do not interfere

- Dynamic: automated growth & shrink dependant on load and budget.

- Separate log files, separate systems

- By Platform Manage, extend this use model to base machine installs

5/19/2008   34

- # On demand multi-site extension
  - Transparent workload transfer to remote sites
 - Integration of resource providers like Amazon, IBM, HP, or my own company sites

My Company data center

Workload transfer

My Resource Provider
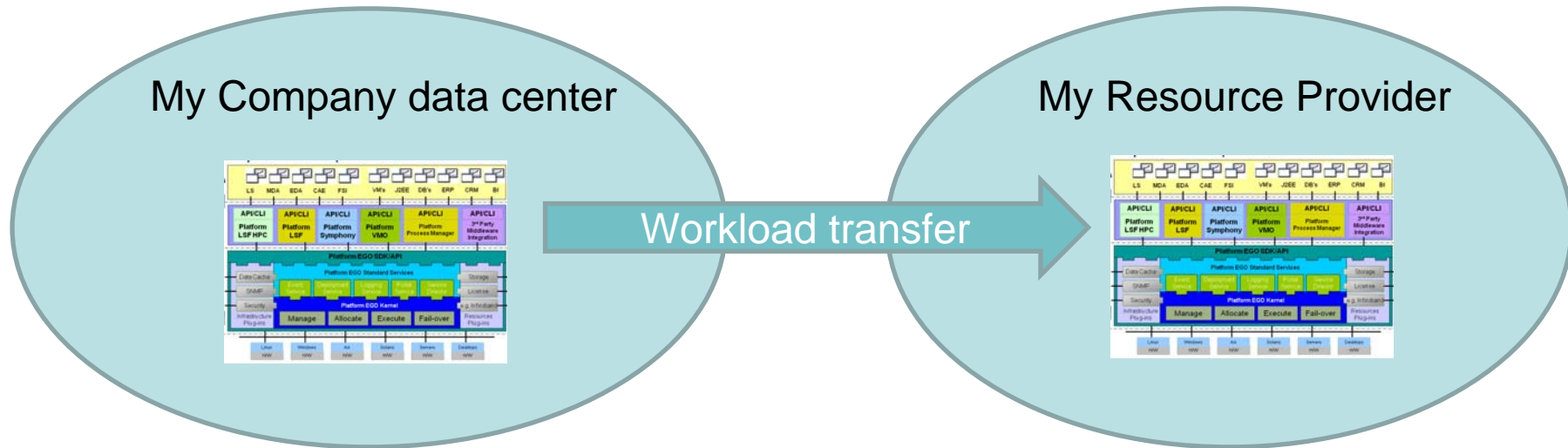
  - Virtualized access to geographically distributed resources
  - In use since on global scale since 1996
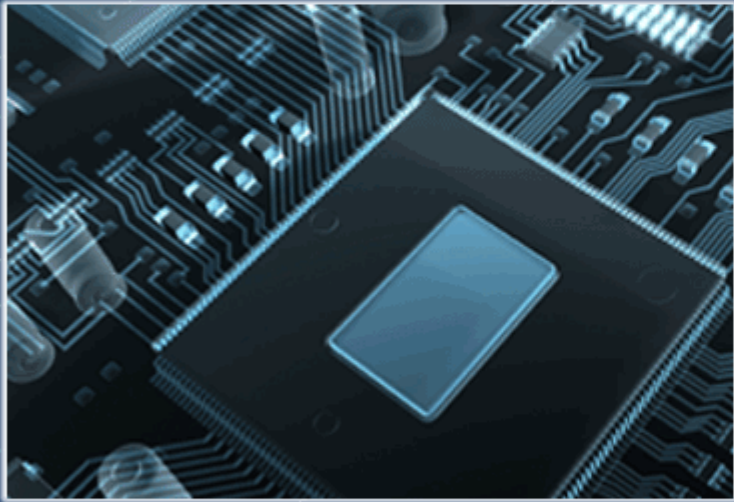
Summary

- VM & Grid together enable significant advantages, cost reductions, and business process speedup
- All components and technologies have proven their value and reliability since many years in industry
- Lets reconnect and detail your solution!
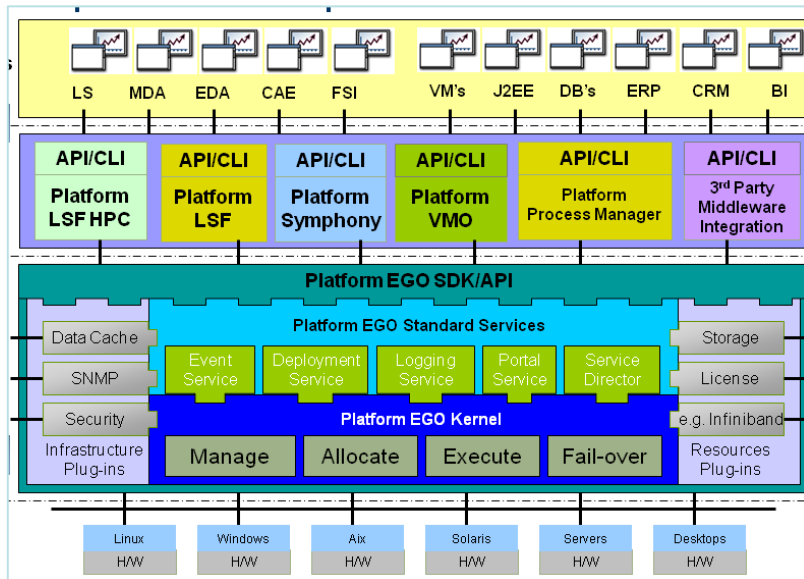
Bernhard Schott
Dipl. Phys.
Business Development Program Manager

Platform Computing GmbH

Frankfurt Office
Direct   +49 (0) 69 348 123 35
Mobile   +49 (0) 171 6915 405
Email: bschott@platform.com
Skype: bernhard_schott
Web: http://www.platform.com/

Platform™

Powering High Performance

Thank you

www.platform.com