

Grid & Virtualization Working Group

OGF23 gridvirt-wg

Erol Bozak, Chair
SAP, Development Architect

June 2008
Barcelona

Wolfgang Reichert, Co-Chair
IBM, Senior Technical Staff Member

OGF IPR Policies Apply



- “I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy.”
- Intellectual Property Notices Note Well: All statements related to the activities of the OGF and addressed to the OGF are subject to all provisions of Appendix B of GFD-C.1, which grants to the OGF and its participants certain licenses and rights in such statements. Such statements include verbal statements in OGF meetings, as well as written and electronic communications made at any time or place, which are addressed to:
 - the OGF plenary session,
 - any OGF working group or portion thereof,
 - the OGF Board of Directors, the GFSG, or any member thereof on behalf of the OGF,
 - the ADCOM, or any member thereof on behalf of the ADCOM,
 - any OGF mailing list, including any group list, or any other list functioning under OGF auspices,
 - the OGF Editor or the document authoring and review process
- Statements made outside of a OGF meeting, mailing list or other function, that are clearly not intended to be input to an OGF activity, group or function, are not subject to these provisions.
- Excerpt from Appendix B of GFD-C.1: “Where the OGF knows of rights, or claimed rights, the OGF secretariat shall attempt to obtain from the claimant of such rights, a written assurance that upon approval by the GFSG of the relevant OGF document(s), any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specification(s) under openly specified, reasonable, non-discriminatory terms. The working group or research group proposing the use of the technology with respect to which the proprietary rights are claimed may assist the OGF secretariat in this effort. The results of this procedure shall not affect advancement of document, except that the GFSG may defer approval where a delay may facilitate the obtaining of such assurances. The results will, however, be recorded by the OGF Secretariat, and made available. The GFSG may also direct that a summary of the results be included in any GFD published containing the specification.”
- OGF Intellectual Property Policies are adapted from the IETF Intellectual Property Policies that support the Internet Standards Process.

Agenda



5 min	Recap of Previous Sessions	Erol
30 min	SAP Scenarios & Demo	Erol
30 min	Interface Requirements & Proposals	Wolfgang
10 min	Relation to RESERVOIR Project	Wolfgang
15 min	Next Steps & Discussion	All

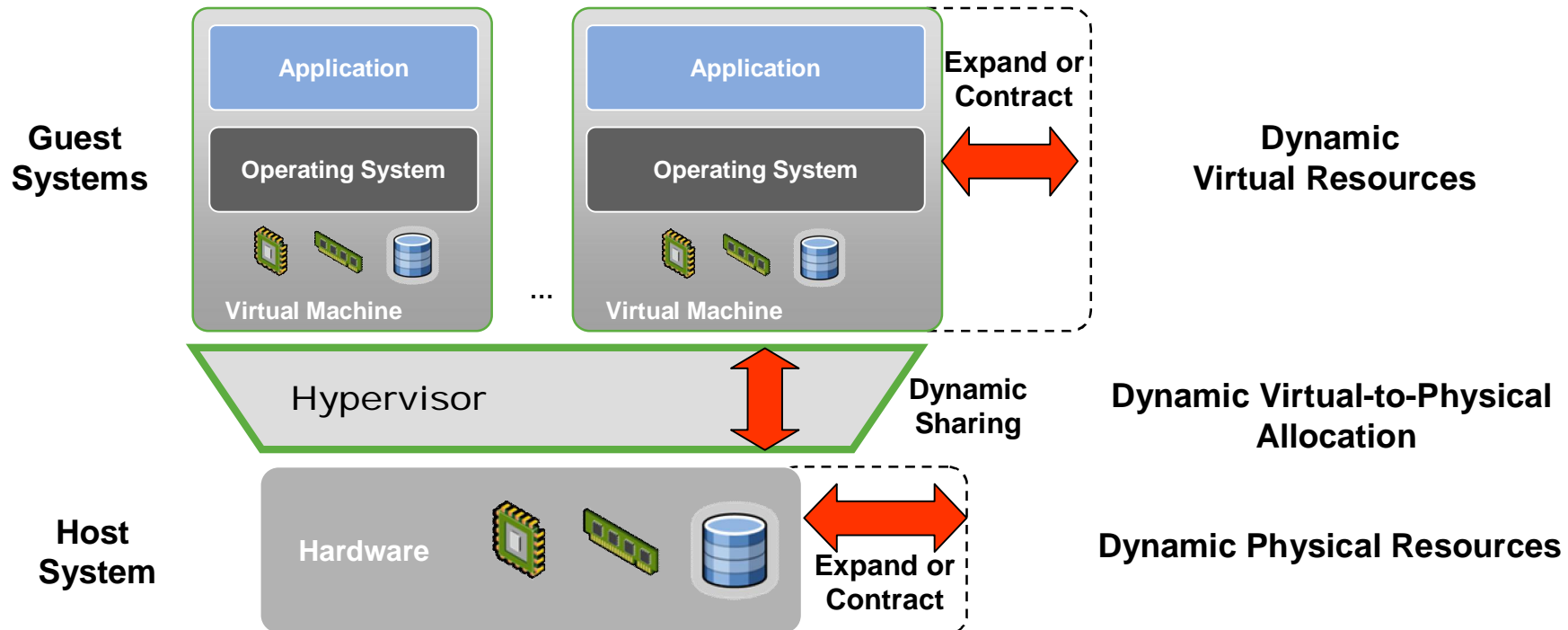
- Recap of Previous Sessions

Goals of the WG



1. Verification that within existing Grid standards the specifications are neutral to virtualized systems and resources
2. Explore how virtualization technologies can be exploited to better support Grid use cases
 - Describe the use cases and scenarios
3. Define the requirements to the Grid architecture for integration with system virtualization platforms
4. Define profiles (i.e. usage recommendations)
 - Align with existing standards, e.g. DMTF - System Virtualization, Partitioning and Clustering (SVPC)

Recap: Virtualization Concept



- Virtualization decouples presentation of resources to consumers (applications) from actual resources through a virtualization layer (Hypervisor)
- Several virtual machines (VMs) may run on a single physical host
- Each VM has its own installed operating system and applications

Recap:

Key Capabilities of Virtualization



- **Creation of virtual systems on-demand**
 - Specify the environment the application / jobs needs to run
 - The environment of the allocating can be pre-configured and persisted as images that can be activated on creation (multiple times if necessary)
- **Dynamic resizing**
 - Change the configuration of virtual system
- **Isolation**
 - Applications / jobs can run isolated from each other
- **Snapshotting**
 - Suspending the virtual system and persisting the state which can be reactivated again
- **Migration**
 - Movement of virtual system among host systems (physical systems)

Recap: Use Cases Overview



Virtualization use cases

- Power saving
- Planned maintenance
- Changing capacity requirements
- Changing capacity offering/availability
- Stateful cloning
- Protecting long running jobs from system failures
- Reproducing situations
- Metering of job resource consumption
- Resource consumption enforcement
- Protection against malware
- Ensuring Security
- Avoiding conflicts
- Emulating an environment for legacy jobs

Virtualization capabilities

Live migration

Dynamic resizing

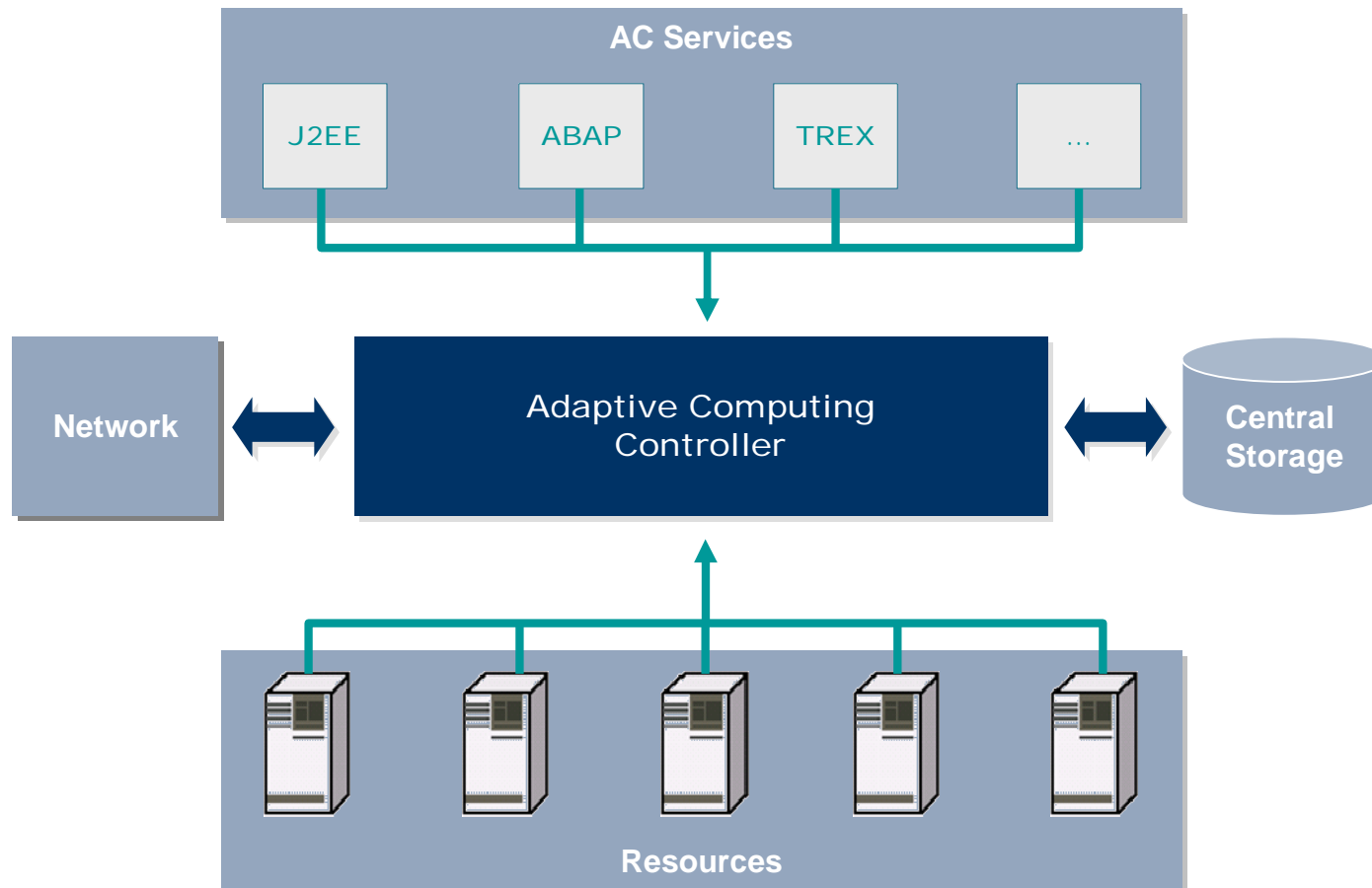
Snapshotting

Isolation

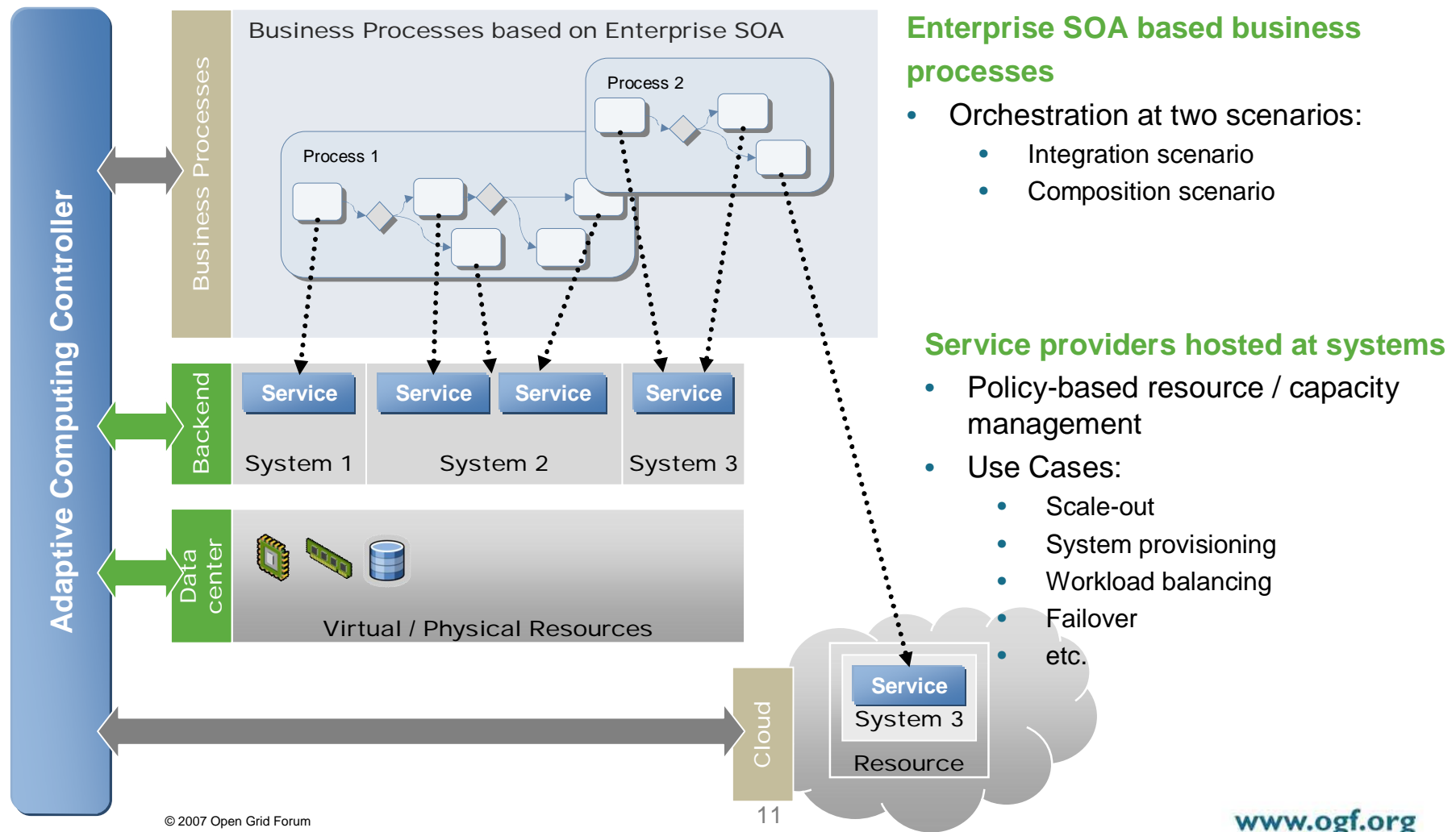
Provisioning

- Adaptive Computing @ SAP
- Demo

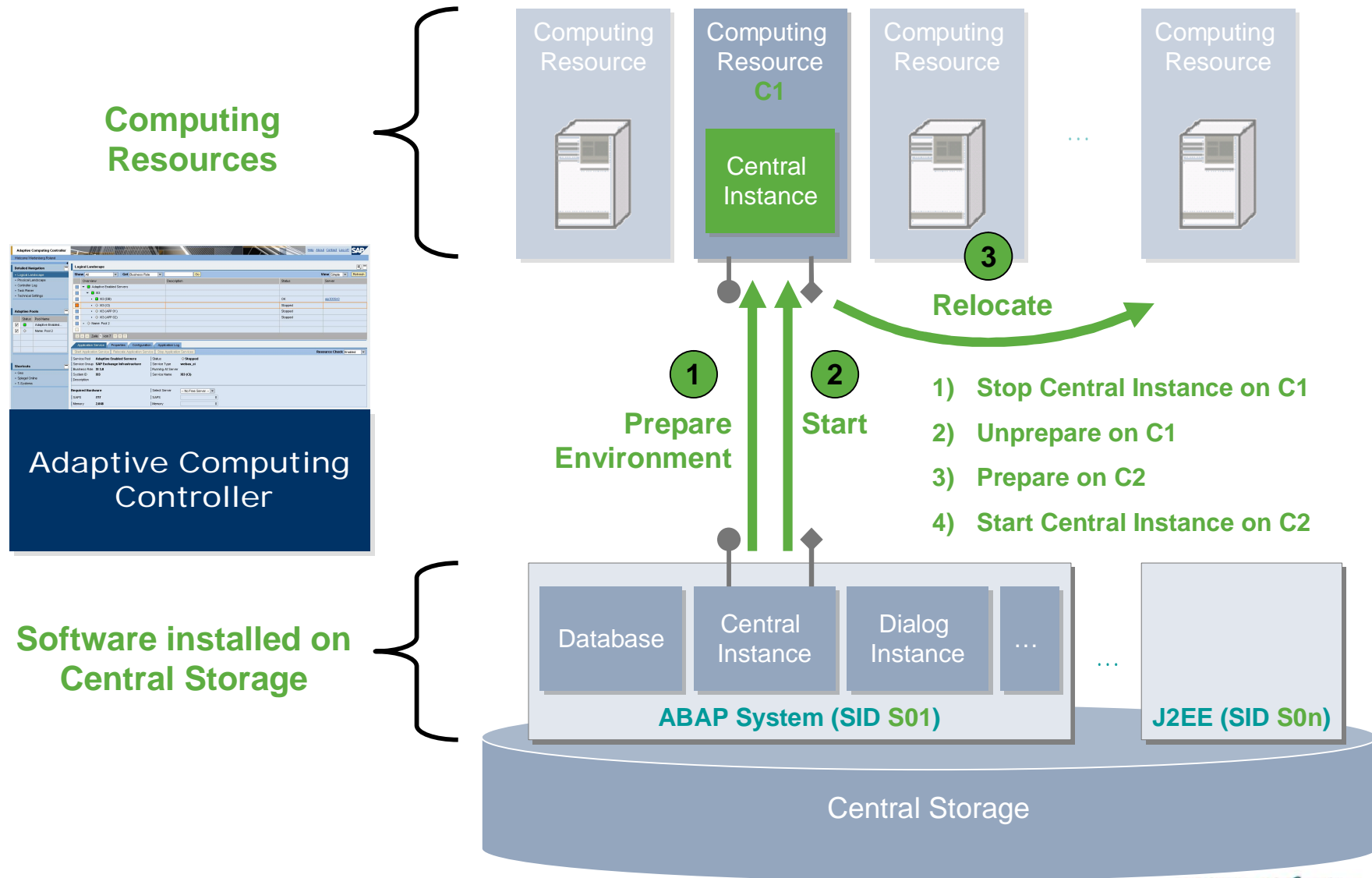
Adaptive Computing @ SAP



Enterprise SOA Resource Management

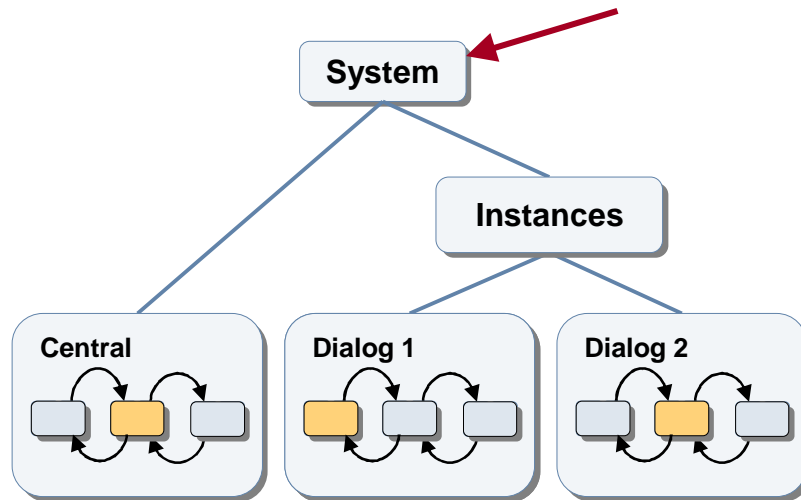


How does Adaptive Computing Work?



Execution Plan Generation

ChangeSystemState("Runnig");



1) System Structure

- Hierarchical
- Leaf node represent executable entities (e.g. SAP Instance)
- Leaf nodes have lifecycle states (stopped, started , prepared etc.)
- Intermediate nodes represent aggregation

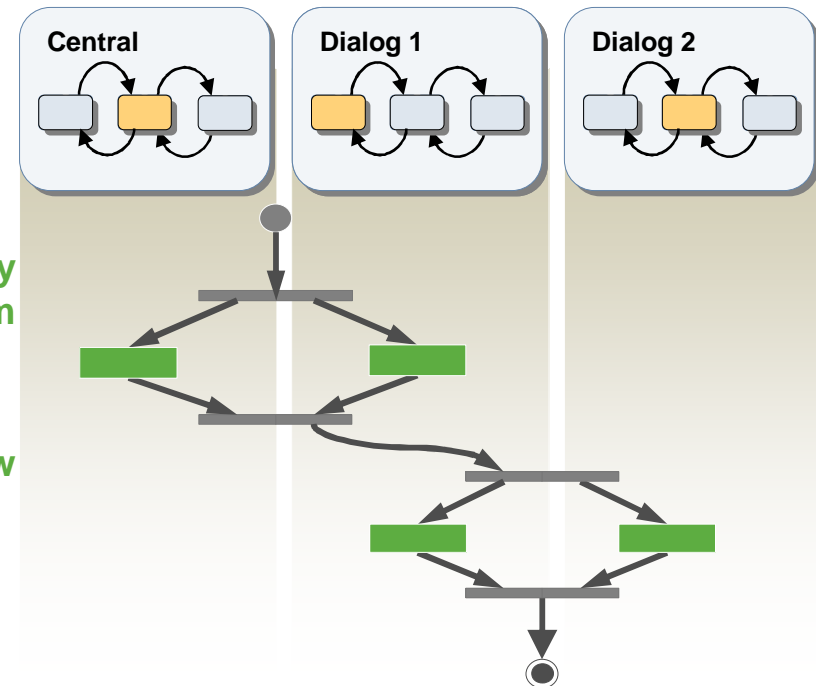
2) Execution Plan

1. Action sequence per node and dependency calculation
2. Execution plan generation (i.e. workflow)

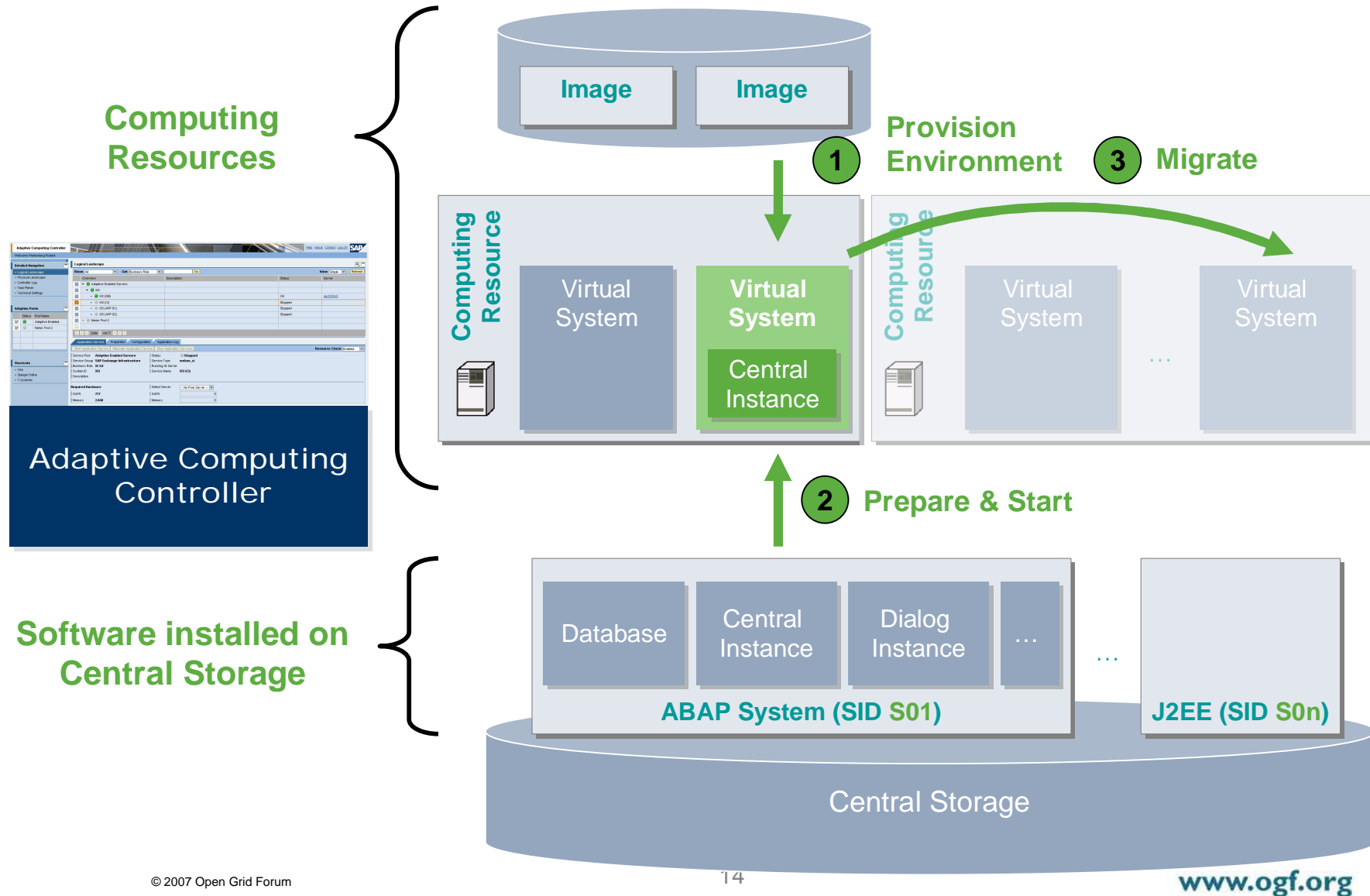
Dependency
Diagram



Workflow



How does Adaptive Computing Work?



Demo



Services

Resources

Virtualization

Logs

Task Planner

Configuration

Related Tasks

Start & Stop Java EE Services

Add a Virtual Machine

ACC Landscape

Refresh

Services

Virtualization: esx003

Group by: All

Auto Refresh (s): 20

Refresh

Find: All

Go

Virtual Machine	OS	CPU Type	Power State	CPU Count	Memory	Hostname	Platform
VMInfrastructure01							VMWare 3.1
esx001	Linux 4.5.5	X86 Power PC	Running	4	6332		VMWare 3.1
vmlx02	Linux 6.1	X86 Power PC	Running	1	2113	HypVsr001	
vmlx03	Linux 6.1	X86 Power PC	Running	2	2338	Hyp098vsr	
vmlx04	Linux 6.1	X86 Power PC	Running	1	1997	Hyp888ght	
esx003	Windows 2000 Server	X86 Power PC	Running	2	6578		VMWare 3.1
vmlx02	Windows XP Service Pack 2	X86 Power PC	Running	1	2334	Hyp890sm	
vmlx03	Linux 4.5	X86 Power PC	Running	1	2567	Hyp768aef	
vmlx02	Linux 4.5	X86 Power PC	Running	1	2145	Hypaae889	

Details

Configuration

Logs

Suspend

Resume

Reset

Power Off

Power On

Migrate

Clone

Snapshot

Revert to Snapshot

View Config

Vendor Mgmt. Interface

Shared Attributes

Guaranteed CPU (mhz): 900

Max. CPU (mhz): 6144

Guaranteed Memory (mb): 2048

Max. Memory (mb): 2048

CPU Utilization (Avg): 70%

Memory Utilization (Avg): 40%

Hostname: HypVsr001

IP Address: 4.4.2

Additional Attributes

Attribute Name	Value	Unit
Attribute Name 01	Linux 6.1	X86 Power PC
Attribute Name 02	Linux 4.5.5	X86 Power PC
Attribute Name 03	Linux 6.1	X86 Power PC

History

Back

Forward

Personalize

Auto Refresh (s): 20

Refresh

Go

Power State	CPU Count	Memory	Hostname	Platform
Running	4	6332		VMWare 3.1
Running	1	2113	HypVsr001	VMWare 3.1
Running	2	2338	Hyp098vsr	
Running	1	1997	Hyp888ght	
Running	2	6578		VMWare 3.1
Running	1	2334	Hyp890sm	
Running	1	2567	Hyp768aef	
Running	1	2145	Hypaae889	

View Config

Vendor Mgmt. Interface

Shared Attributes

CPU speed per core (mhz): 900

Memory Capacity (Mb): 6144

Hypervisor Name: HypVsr001

Hypervisor Version: 4.4.2

Suspend Supported: Yes

Live Migration Supported: No

Max. supported VMs: 20

Max. Running VMs: 10

Additional Attributes

Attribute Name	Value	Unit
Attribute Name 01	Linux 6.1	X86 Power PC
Attribute Name 02	Linux 4.5.5	X86 Power PC
Attribute Name 03	Linux 6.1	X86 Power PC
Attribute Name 04	Linux 4.5.5	X86 Power PC
Attribute Name 05	Linux 6.1	X86 Power PC

CPU & Memory Utilization Report

Total CPU Utilization

Total Memory Utilization

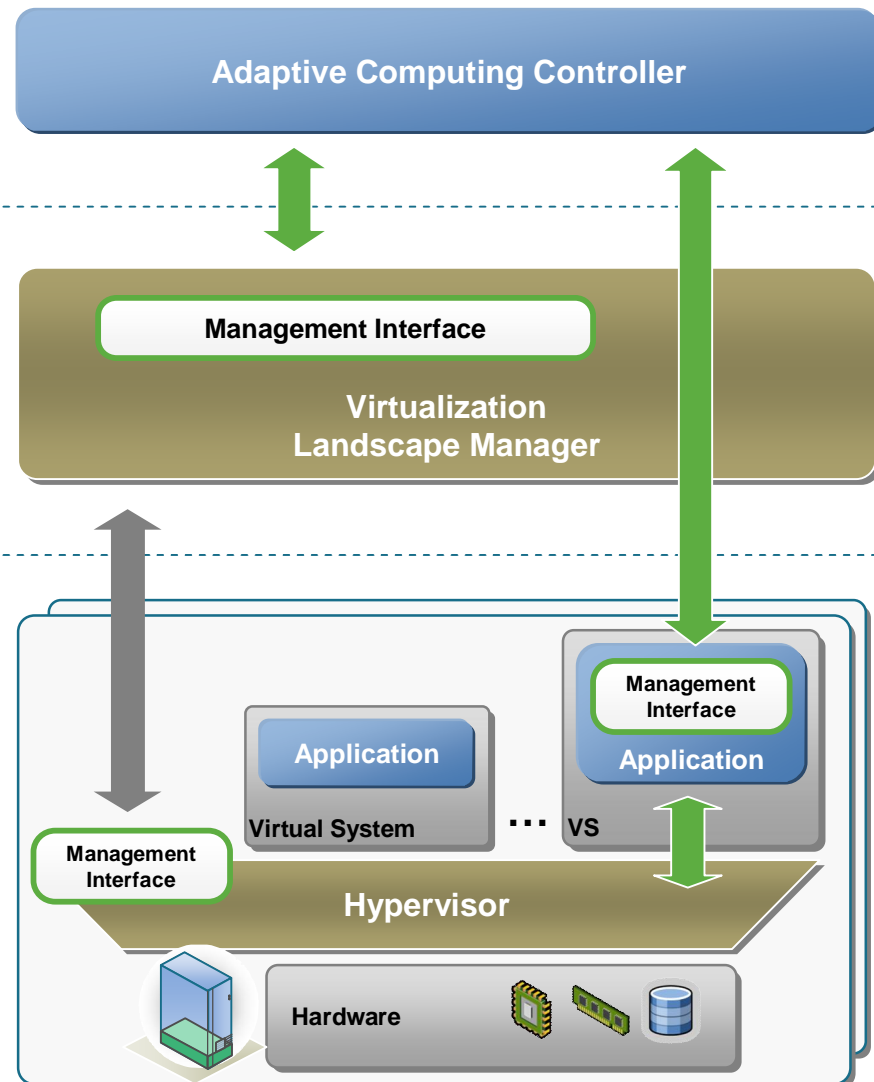
Aggregate Report

Hypervisor

VMs

Not in use

Levels of Integration



© 2007 Open Grid Forum

- Adaptive Computing Controller provides „higher level“ capabilities for managing SAP systems and resources
- Manages physical as well as virtual resources
- Platform specific „Virtualization Landscape Manager“ (VLM) manages distributed sets of virtual systems and resources in a landscape
- Proprietary interfaces exposed at this level !
- Platform specific interface exposed at the „Virtualization Engine“ (VE), e.g. Hypervisor
- Standardization is focused on this level
- Initiative at the DMTF: SVPC (Server Virtualization, Partitioning & Clustering)
- Application awareness needed !

Capabilities



- **Operations**
 - Types:
 - Create (SVPC: Define) (see provisioning)
 - Destroy
 - Power on (SVPC: Activate) / Power off (SVPC: Deactivate) / Reset
 - Shutdown, Reboot
 - Migrate
 - Clone
 - Retrieving “operation logs” of executed and/or executing operations
 - Asynchronous / synchronous processing of operations
 - Tracking progress
 - Cancelable operations
- **Provisioning of Virtual Systems**
 - Templates based virtual system creation
 - OVF (Open Virtual machine Format) based provisioning
- **Configuration**
 - Configuration exclusively for resizing
 - 1. Activation of predefined size configuration (e.g. Small, Medium, Large, XL)
 - 2. Customization via standard metrics

Capabilities

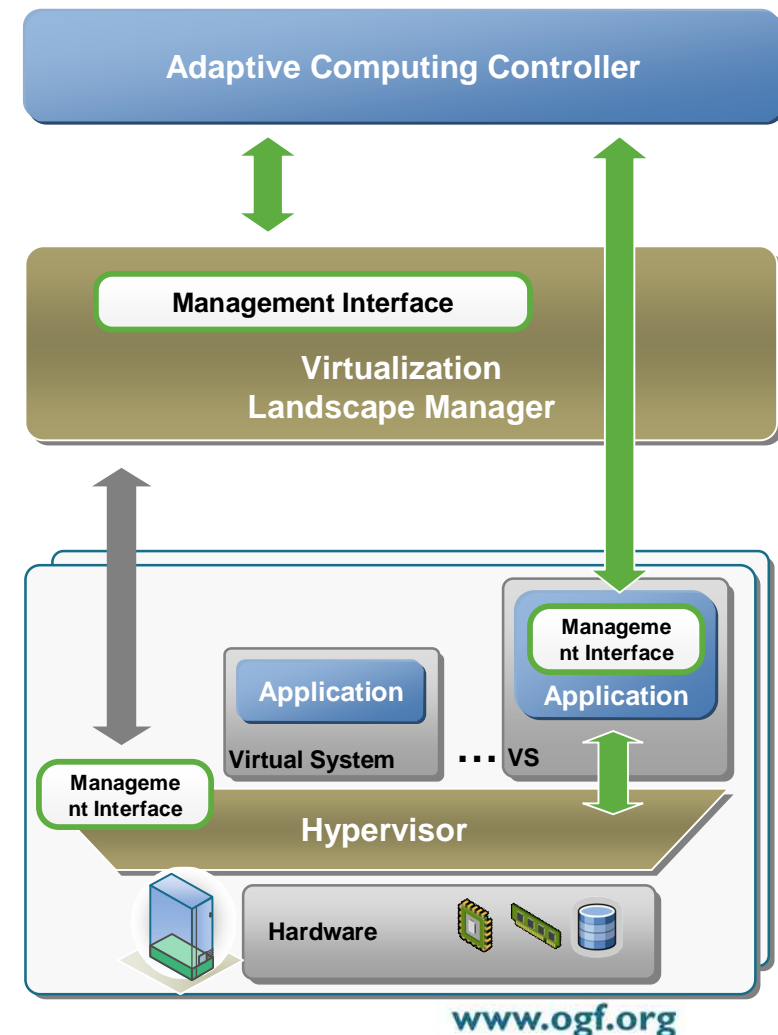


- **Inventory / Discovery / Monitoring**
 - Inventory & discovery of templates and virtual systems and hosts
 - Monitoring of virtual systems and hosts
 - Selective monitoring (through attribute specification) for all managed entities in the virtualization landscape, e.g. virtual systems, host systems, operations (tasks) etc.
- **Snapshot**
 - Creation of snapshots and reverting virtual systems from its snapshots
- **Other**
 - Retrieving alerts from the virtualization platform
 - Alerts may be raised for errors & exceeding thresholds (i.e. policy violation)
 - Alerts have to have a reference to the alert details / situation

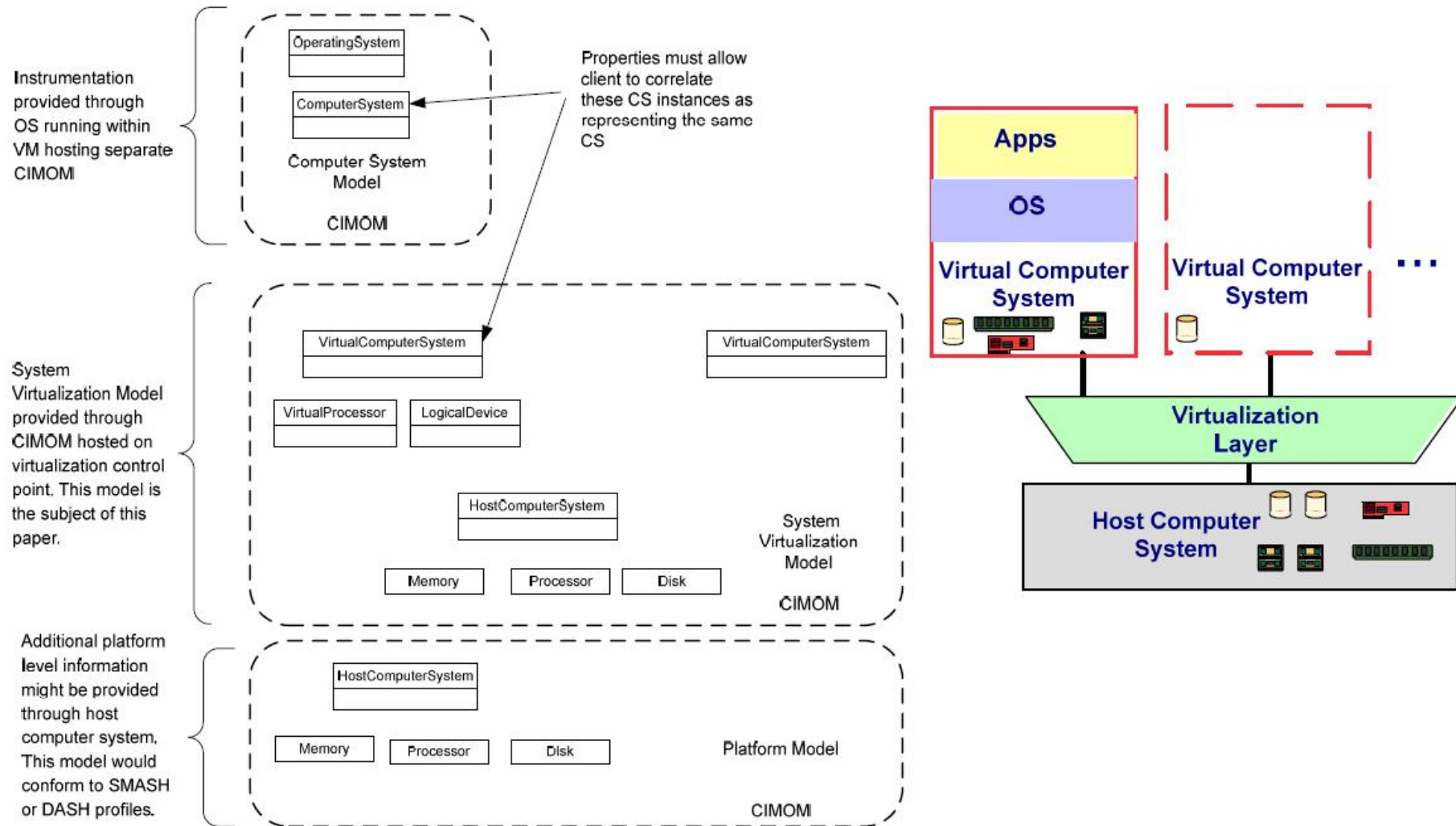
- Interface Requirements & Proposals

Management Stack

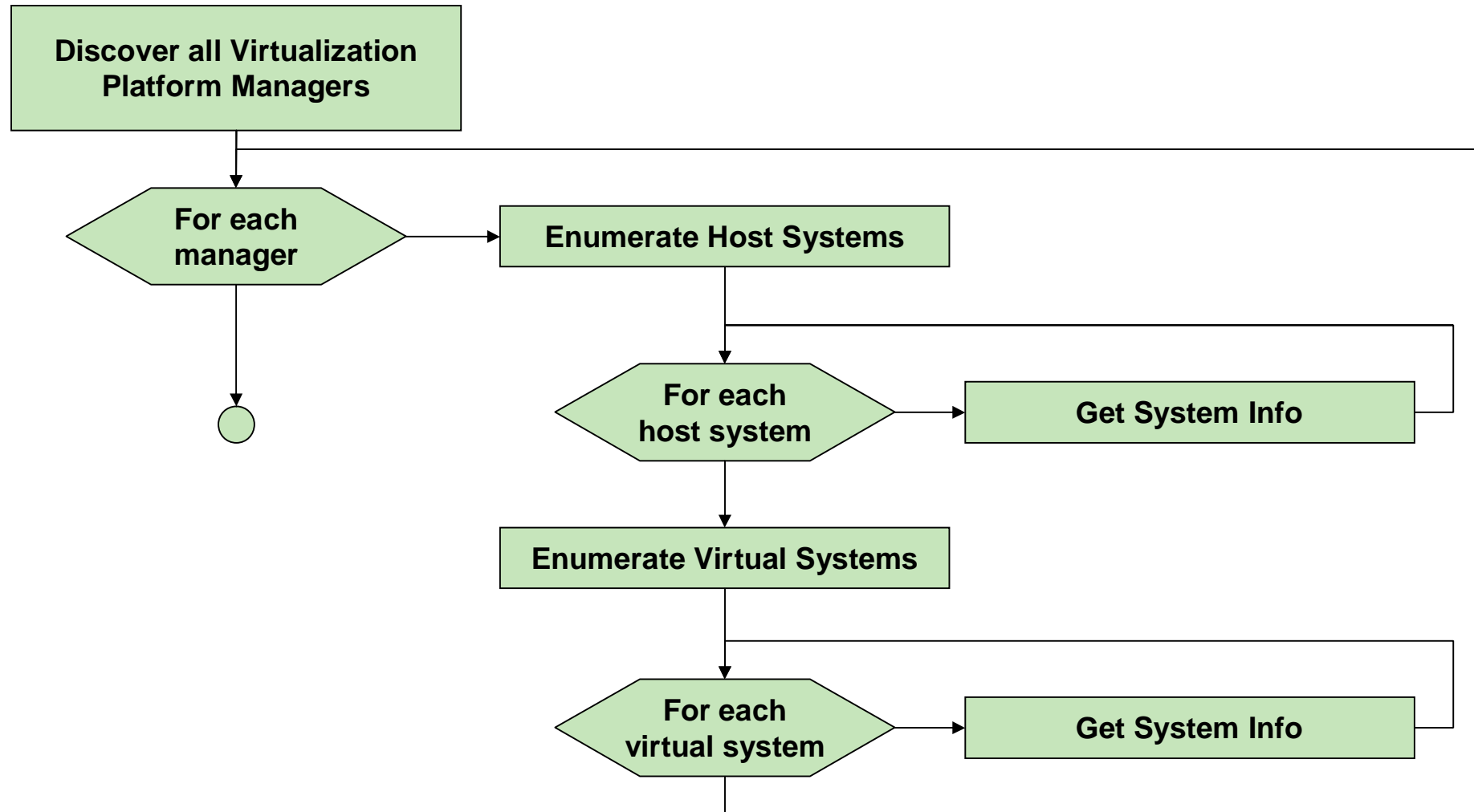
- Cloud / Grid
- Datacenter / Landscape (*heterogeneous environment*)
 - SAP ACC
 - ...
- Cluster (*homogeneous environment*)
 - Virtual Center (VMware)
 - HMC (Hardware Management Console, IBM)
 - IBM Director
 - ZENworks Orchestrator (Novell)
 - ...
- Virtual Systems running OS
- Servers / Hypervisors
 - IBM POWER Hypervisor, z/VM
 - VMware ESX
 - XEN
 - ...



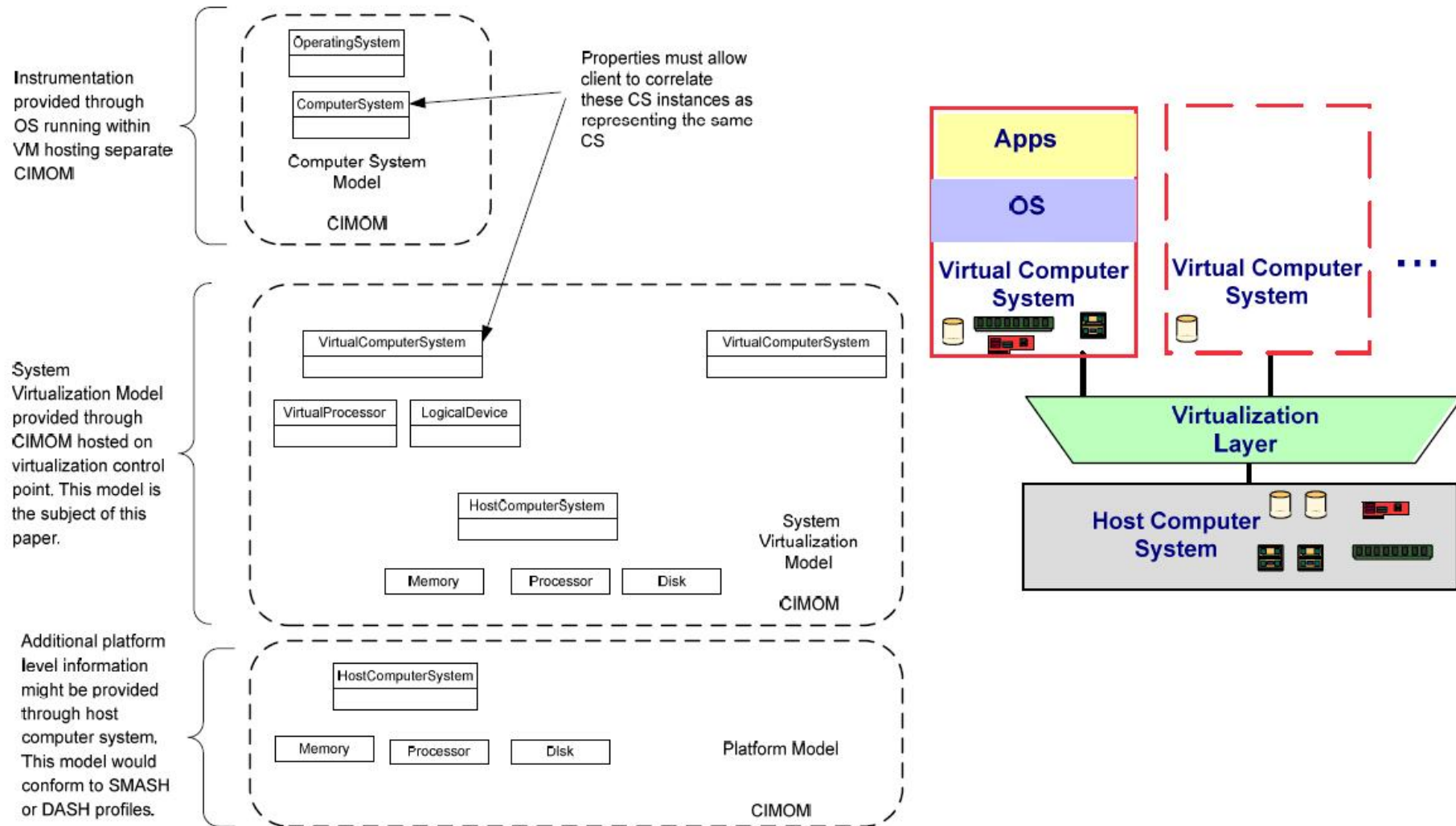
CIM System Virtualization Model



Topology Discovery



Virtualization Modeling Environment



Topology Discovery: CIM Approach



- CIM = distributed model
 - Manager needs to establish contact with each CIMOM
 - SLP broadcast to discover all CIMOMs
 - Query CIM_RegisteredProfile → “System Virtualization”
 - Enumerate instances
 - Virtualization manager knows ComputerSystem data, but not OperatingSystem data
 - Need to contact CIMOMs in OS for application driven scenarios
 - Many (hundreds/thousand) CIM connections required
 - Associators indicate relation
 - CIM ComputerSystem ID is not hostname
- Higher level API could be more efficient than CIM

Topology Discovery...



GetAllHostSystems

IN

- With Filter: Hardware architecture (INTEL, POWER, ...)
- With Filter: Virtualization type (XEN, VMware, z/VM, ...)

optional

optional

OUT

- List of system identifiers

GetAllVirtualSystems

IN

- With Filter: State (Active, Defined, ...)
- With Filter: RunningOperatingSystem (Linux, Windows, AIX ...)

optional

optional

OUT

- List of system identifiers

GetAllVirtualSystemsOnHost

IN

- System identifier
- With Filter: State (Active, Defined, ...)
- With Filter: RunningOperatingSystem (Linux, Windows, AIX ...)

optional

optional

OUT

- List of system identifiers

Topology Discovery...



GetSystemInfo

IN

- List of system identifiers

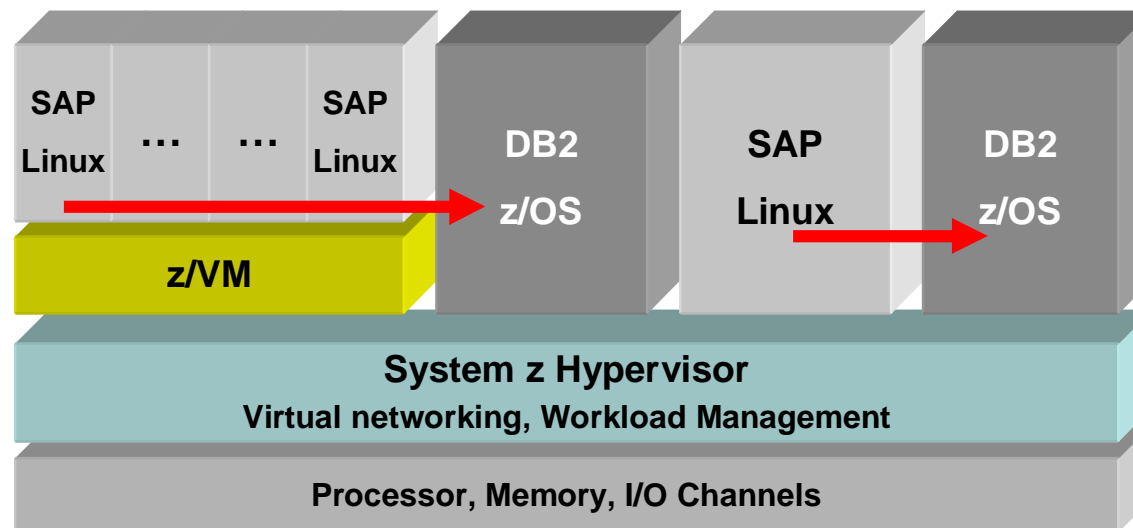
OUT

- List of base properties for each system

Topology Discovery: Implications



- Virtualization is not limited to two layers, rather it is a tree
 - A virtual system might be a host system for other virtual systems
 - The tree is navigated starting with the (physical =) lowest level host system



Multiple virtualization layers on IBM System z

Virtual System Settings and Capabilities



GetVirtualSystemSettings

IN

- System identifier

OUT

- List of settings

GetResourceAllocationSettings

IN

- System identifier

OUT

- List of resource allocation data
- Available resources in the pool

GetVirtualSystemManagementCapabilities

IN

- System identifier

OUT

- List of supported management services

ChangeVirtualSystemSettings

IN

- System identifier
- List of changed settings

OUT

- List of settings

ChangeResourceAllocationSettings

IN

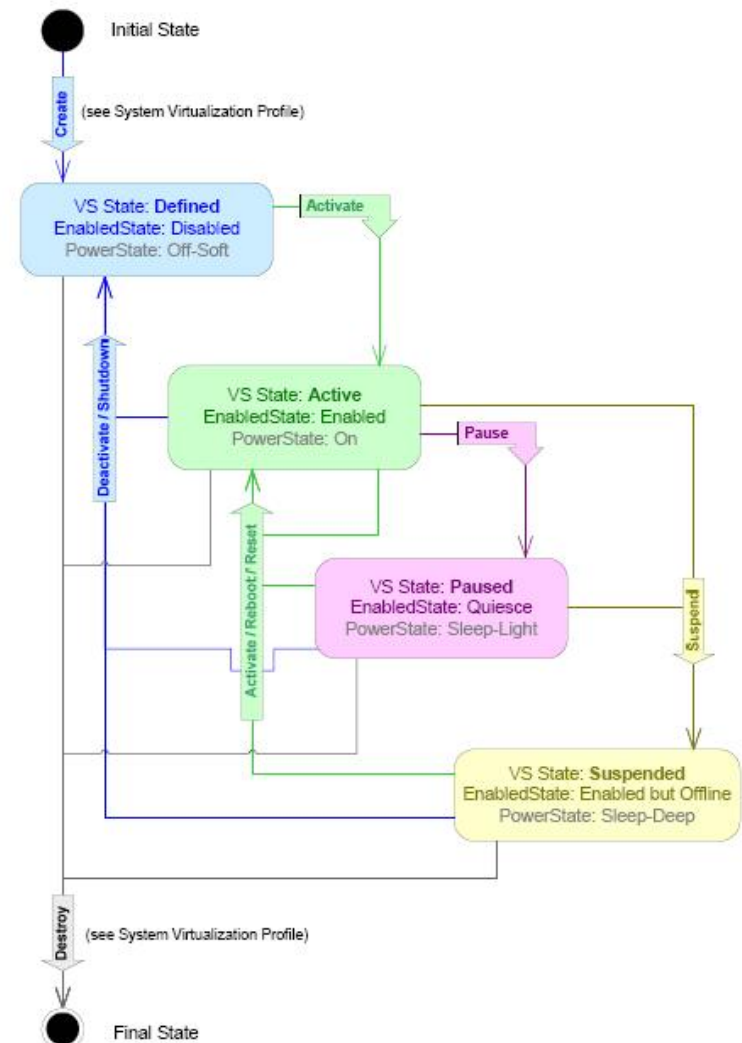
- System identifier
- List of changed resource allocation data

OUT

- List of resource allocation data

Virtual System Lifecycle Operations

- Virtual system state transitions
 - Create / Destroy
 - Activate / Deactivate
 - Pause / Suspend
 - Shutdown / Reboot / Reset
- Value-added operations
 - Migrate
 - Clone
 - Create/Apply/Destroy snapshot
- Helper services
 - Query progress of asynchronous operations
 - Cancel asynchronous operations
 - Retrieve operation logs



CIM virtual system state model

Snapshot / Live Migration



CreateSnapshot

IN

- Virtual System identifier
- Name
- Description

OUT

- Snapshot identifier

ApplySnapshot

IN

- Virtual System identifier
- Snapshot identifier

DestroySnapshot

IN

- Snapshot identifier

ListAllSnapshots

IN

- Virtual System identifier

OUT

- List of snapshot identifiers and attributes

MigrateVirtualSystem

IN

- Virtual System identifier
- Target host system identifier
- Mode (Test / Prepare / Execute)
- Properties [*optional*]

OUT

- Job identifier

- Relation to RESERVOIR project

RESERVOIR



<http://www.reservoir-fp7.eu/>



- Home
- Research
- Consortium
- Press
- Presentations
- Publications
- Project Deliverables
- Contact Us

Welcome to Reservoir!


Resources and Services Virtualization without Barriers

Resources and Services Virtualization without Barriers is an European Union FP7 funded project that will enable massive scale deployment and management of complex IT services across different administrative domains, IT platforms and geographies. The project will provide a foundation for a service-based online economy, where - using virtualization technologies - resources and services are transparently provisioned and managed on an on-demand basis at competitive costs with high quality of service.

Web 2.0 is rapidly taking hold, offering "the web as a platform". In parallel, traditional client-server computing is starting to lose ground as a new paradigm emerges - the Cloud Computing paradigm. Cloud Computing allows data centers to operate more like the Internet by enabling computing across a distributed, globally accessible fabric of resources, delivering service based on demand over the web, reducing software complexity and costs, expediting time-to-market, improving reliability and enhancing accessibility of consumers to government and business services. Thus, Cloud Computing represents a true materialization of Service-Oriented Computing's visionary promise. In RESERVOIR, we are developing breakthrough system and service technologies that will serve as the infrastructure for Cloud Computing. We aim to achieve this goal by creative coupling of virtualization, grid computing, and business service management techniques.

RESERVOIR: Consortium




The RESERVOIR logo, with the word "RESERVOIR" in blue capital letters and a stylized orange grid pattern to the right.

Consortium


- Home
- Research
- Consortium**
- Press
- Presentations
- Publications
- Project Deliverables
- Contact Us

Resources and Services Virtualization without Barriers


Partners in the consortium come from both industry and academia, and represent:

The IBM logo, consisting of the letters "IBM" in a blue, horizontally-striped font.


[IBM Haifa Research Lab](#) - HRL is recognized as a center of competence in system technology and virtualization in particular, and is involved in advanced research and development in these areas. A recent achievement of HRL is a key contribution to the development of the IBM Virtualization Manager product. HRL is the Project Coordinators for RESERVOIR

The Telefonica logo, with the word "Telefonica" in yellow script on a dark blue background.


[Telefónica Investigación y Desarrollo](#) - the research arm of the Telefónica Group, which is one of the world's largest telecom companies

The UCL logo, with the letters "UCL" in white on a black background, and a small white building icon to the left.

[University College of London](#) - One of the premier universities in the United Kingdom with a strong emphasis on research

The Umeå University logo, featuring a blue circular emblem with a white building and the text "Umeå University" around it.

[Umeå University](#) - The largest and oldest university in northern Sweden with over 27,000 students

The SAP logo, with the letters "SAP" in white on a blue background.

[SAP Research](#) - The technology research department of SAP, the world's leading provider of e-business software solutions

RESERVOIR: SAP Use Cases and Requirements



- Creating a manifest of a service application
 - Images, contextualization scripts, DB content, all other configuration data
 - Means to contextualize and customize
- Provisioning a service application from a manifest
- Dynamic adjustment of resource allocation
 - Capacity planning
 - Automatic adaptive resource allocation / Self-optimization based on SLA and actual workload
- Elastic array of virtual execution environments
 - Dynamic scale-out by adding virtual servers to a service application
- Live migration
- . . .

- Next Steps & Discussion

Next Steps



1. Publish work group paper for public comments
 - Document type: “Informational”
2. Alignment with DMTF System Virtualization Working Group
3. Alignment with RESERVOIR