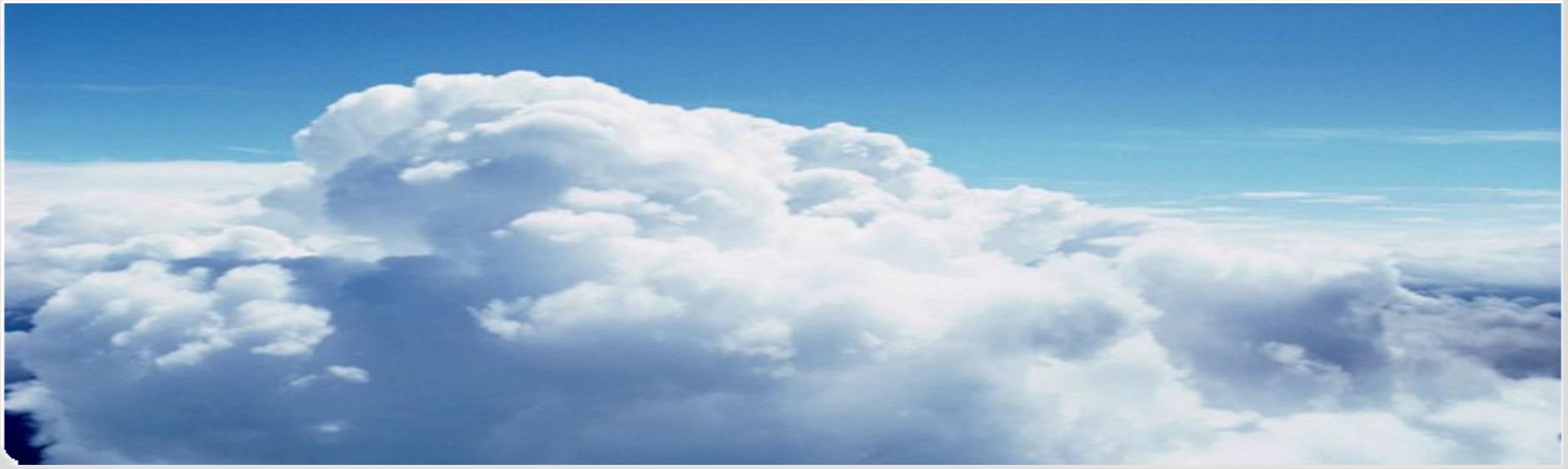
Open Cirrus: Towards an Open Source Cloud Stack

Karlsruhe Institute of Technology (KIT)
HPC2010, Cetraro, June 2010

Marcel Kunze



Open Cirrus Cloud Computing Research Testbed

http://opencirrus.org

- **An open, internet-scale global testbed for cloud computing research**
  - A tool for collaborative research
  - Data center management & cloud services
- **Resources**
  - Multi-continent, multi-datacenter, cloud computing system
  - Federated “Centers of Excellence” around the globe
    - each with 100–400+ nodes and up to ~PB storage
    - and running a suite of cloud services
- **Structure**
  - Sponsors: HP Labs, Intel Research, Yahoo!
  - Partners: CMU, ETRI, IDA Singapore, KIT, MIMOS, RAS, UIUC
  - Discussion with: CESGA, UNIDO, China Mobile
  - Cloud system and application development
- **Cover story in IEEE Computer, vol 43, no 4, April 2010**

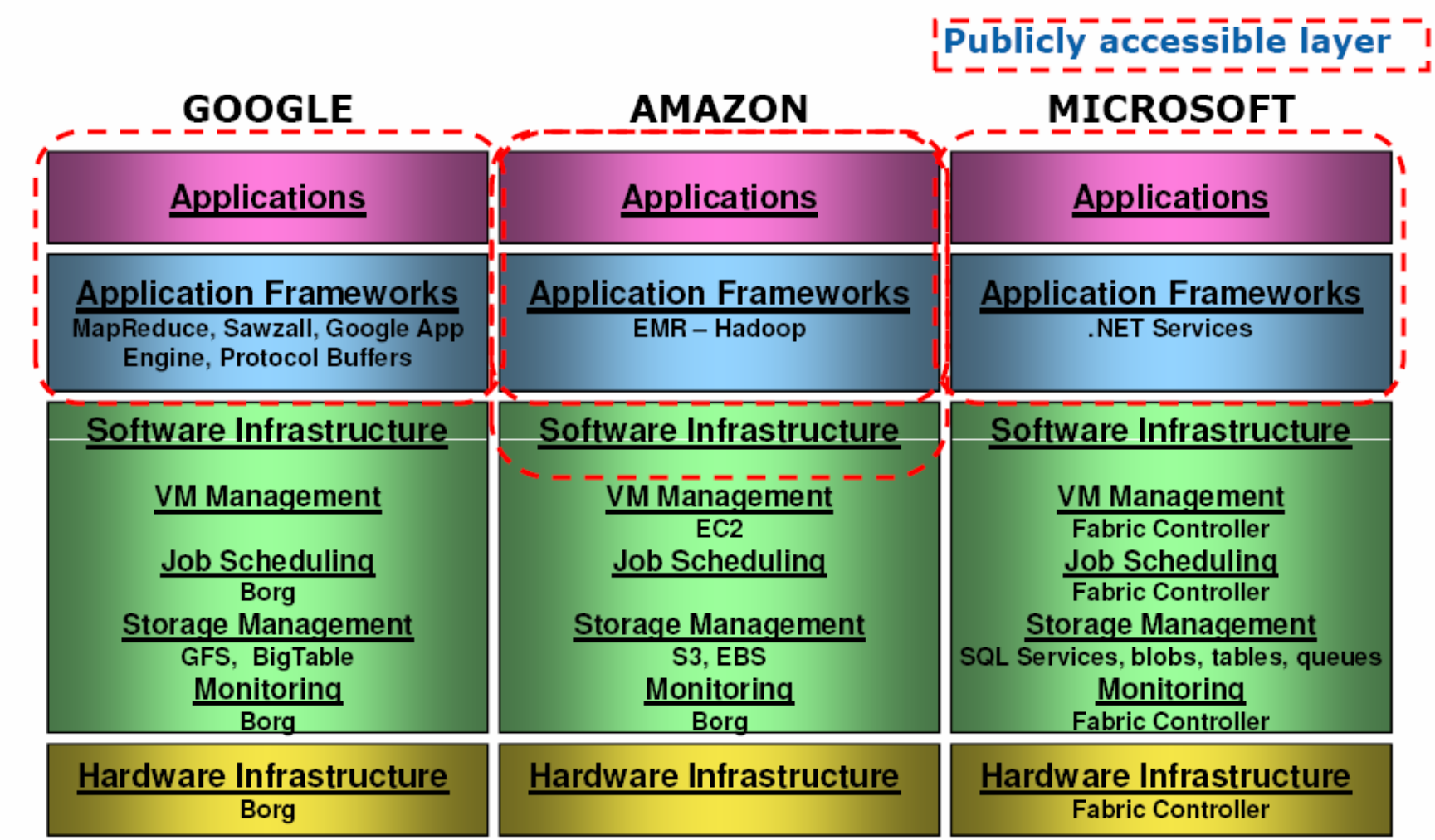
Cloud Systems Research

- **Open Cirrus is seeking research in the following areas:**
  - Cloud services
  - Datacenter federation
  - Datacenter management and automation
  - Data-intensive applications and systems

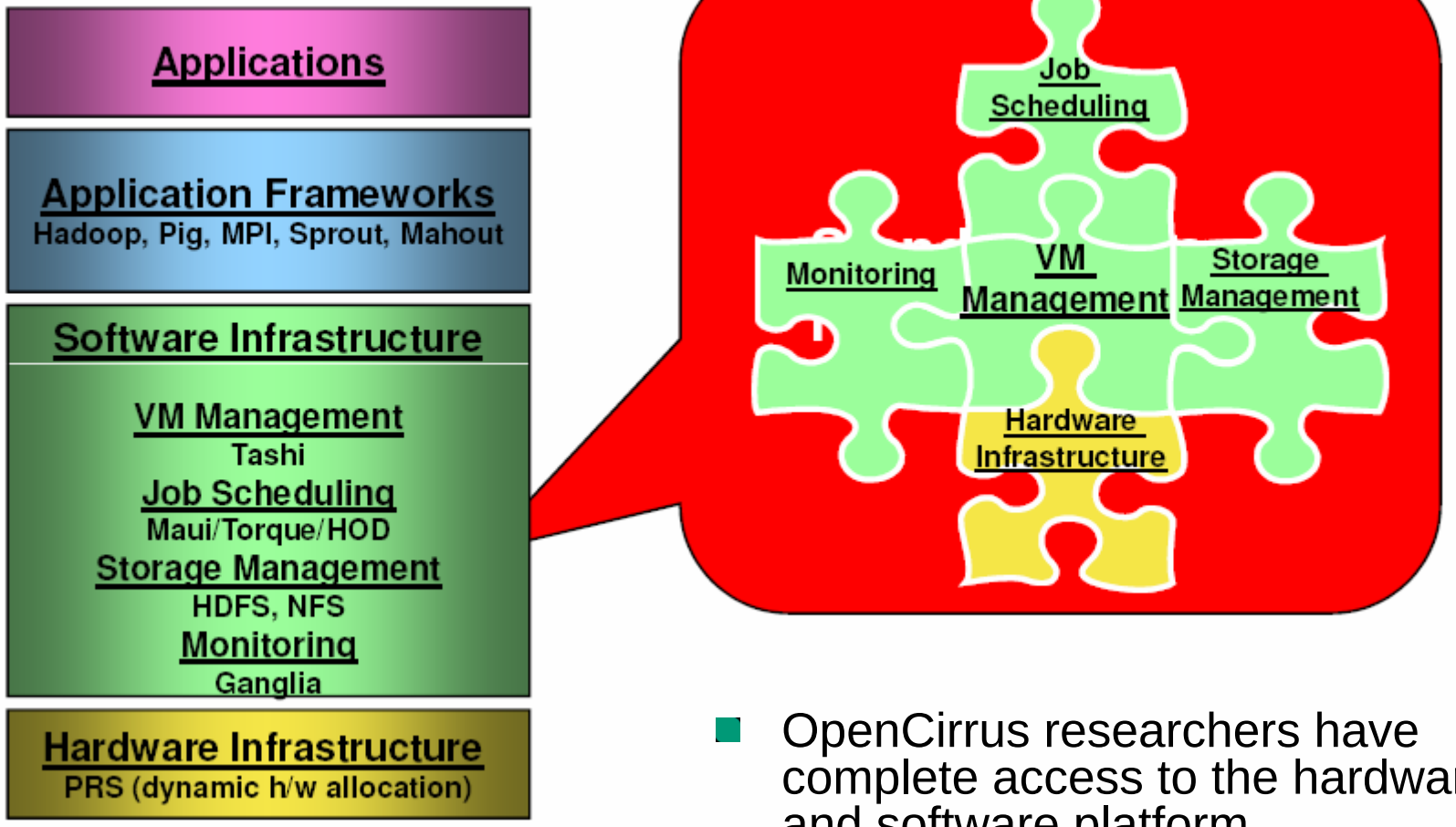
- **Research requirements**
  - Perform experiments also on a low system level
  - Flexible cloud computing framework
  - Compare different methodologies and implementations

- **Simple, transparent, controllable cloud computing infrastructure**

Proprietary Cloud Computing Stacks

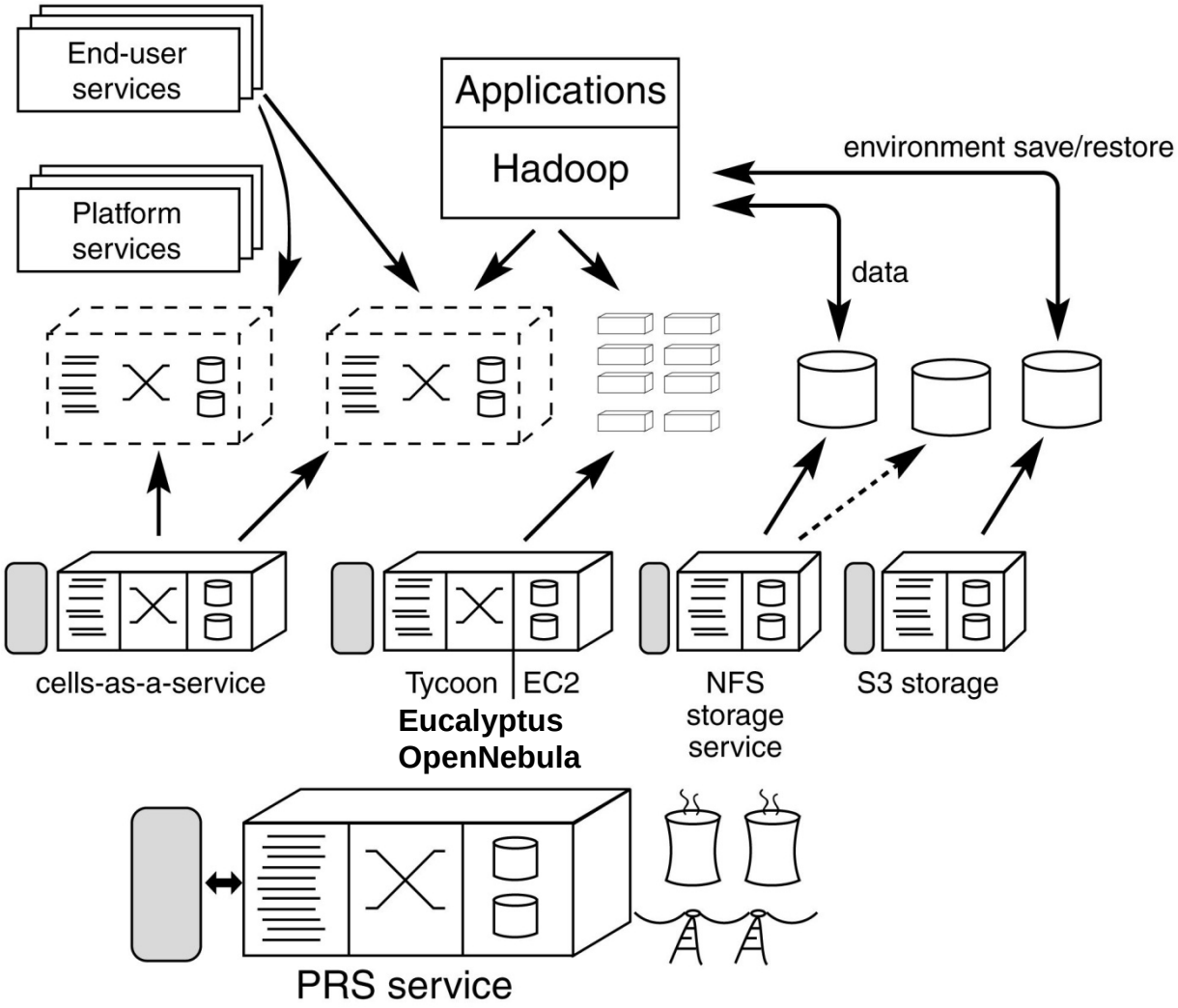


Open SourceCloud Stack



- OpenCirrus researchers have complete access to the hardware and software platform

Open Cirrus Blueprint



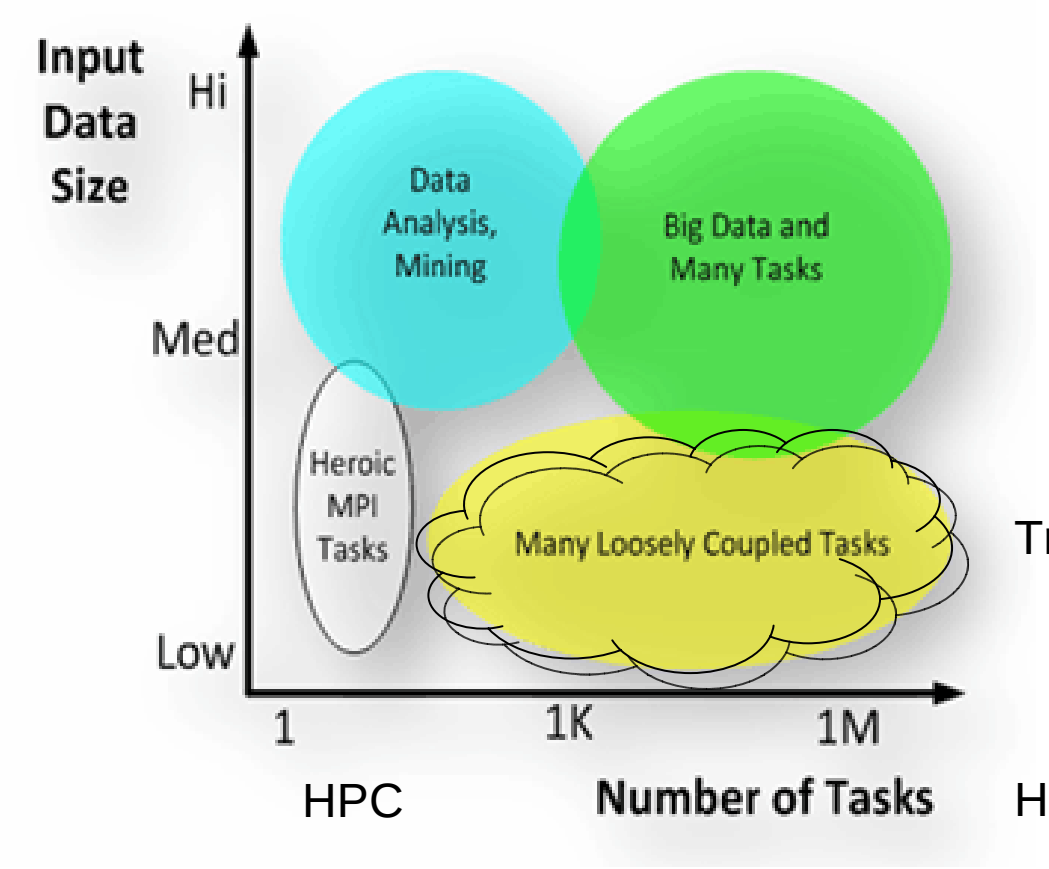
Cloud Application Services

Virtual Resource Sets

Cloud Infrastructure Services

IT-Infrastructure Layer (Physical Resource Sets)

HPC vs. HTC vs. MTC (Many Task Computing)

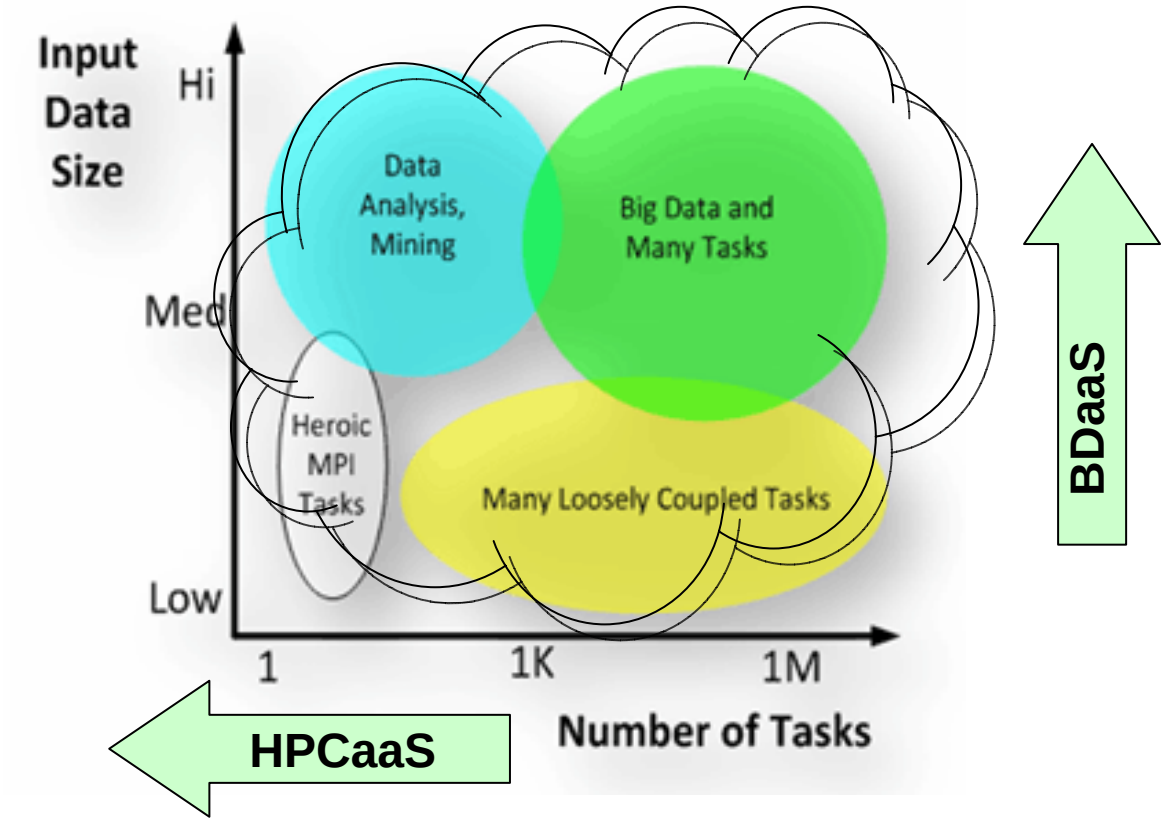


MTC

Traditional Cloud

Source: I.Foster

Extension of the Cloud Space

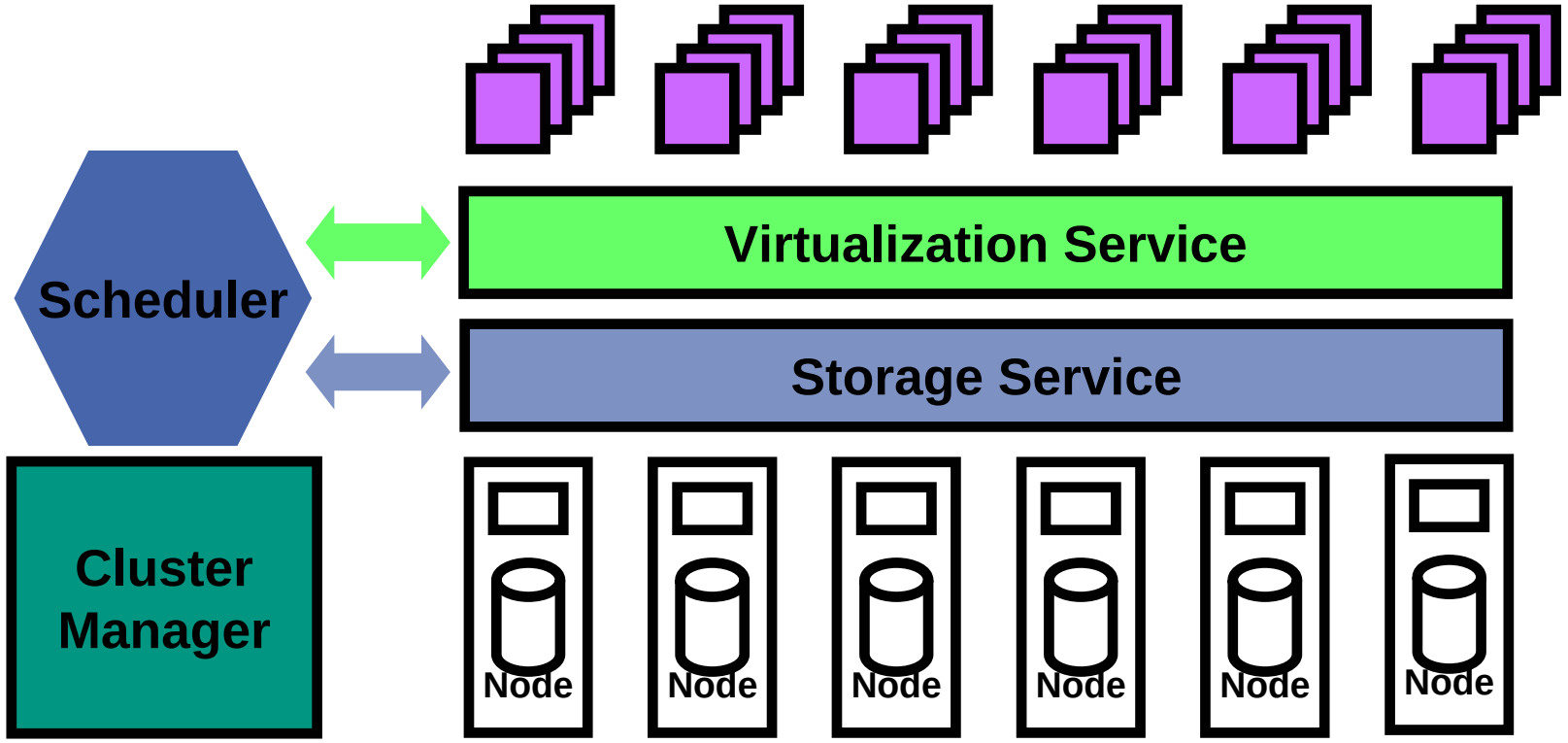


- **Cloud development towards HPC and Big Data as a Service**

Big Data as a Service: Tashi

http://wiki.apache.org/incubator/TashiProposal

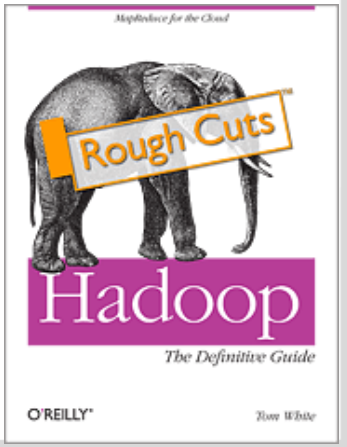
OpenCirrus™ at Intel Research Pittsburgh



- Virtualization of CPU and storage
- Co-Scheduling of CPU and data

Big Data as a Service: Hadoop

- Reproduce the proprietary software infrastructure developed by Google as Open Source
- Hadoop implements
  - Parallel programming model (MapReduce)
  - Hadoop Distributed File System (HDFS)
  - Parallel database (HBase)
  - Programming environment (Pig)
  - Data warehouse infrastructure (HIVE)
  - Data collection and analysis (Chukwa)
  - Machine based learning (Mahout)
- Largest Cluster at Yahoo!: 32.000 cores and 16 PetaByte storage (1PB/16h)

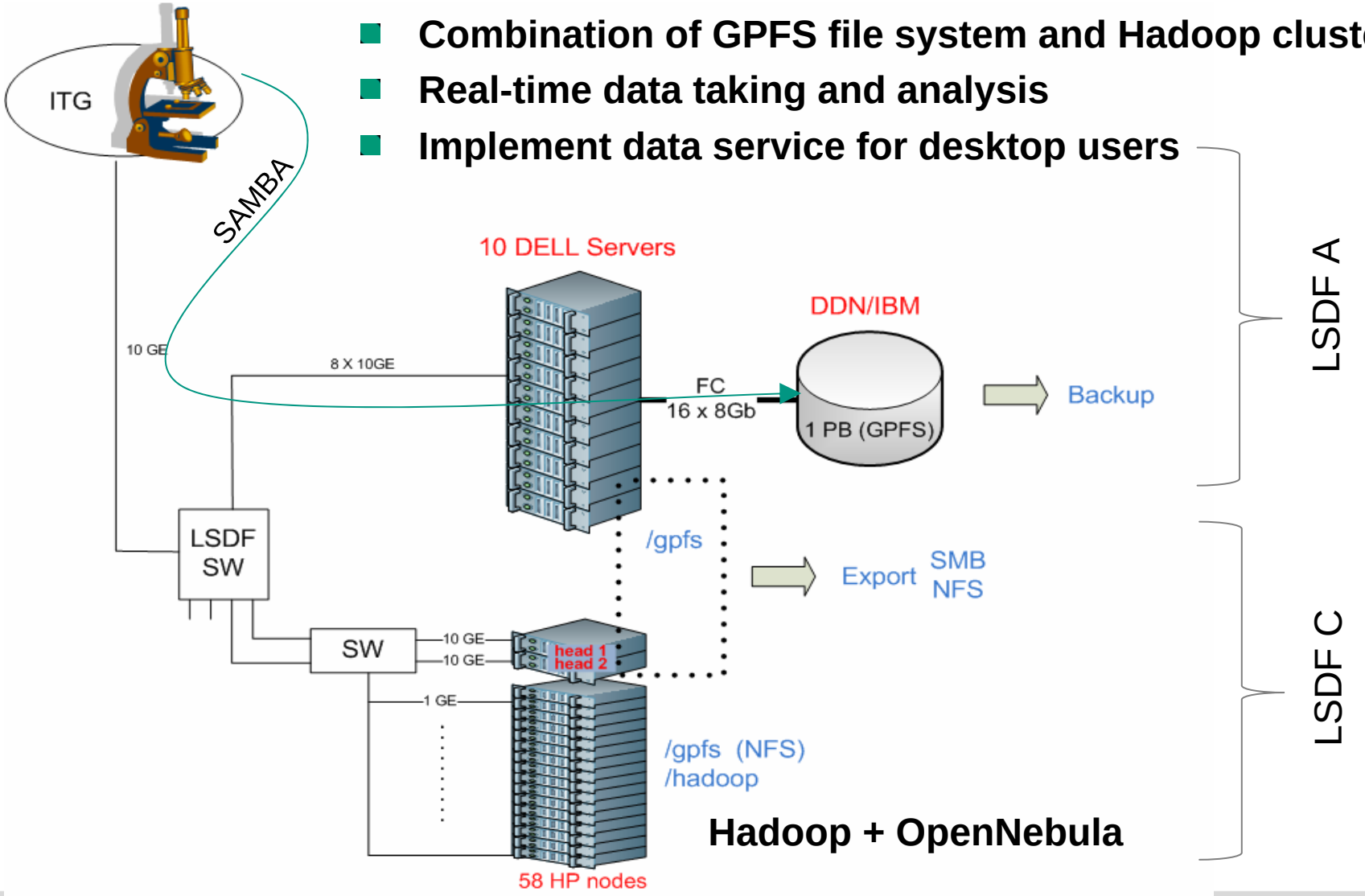


Hadoop Cluster at KIT

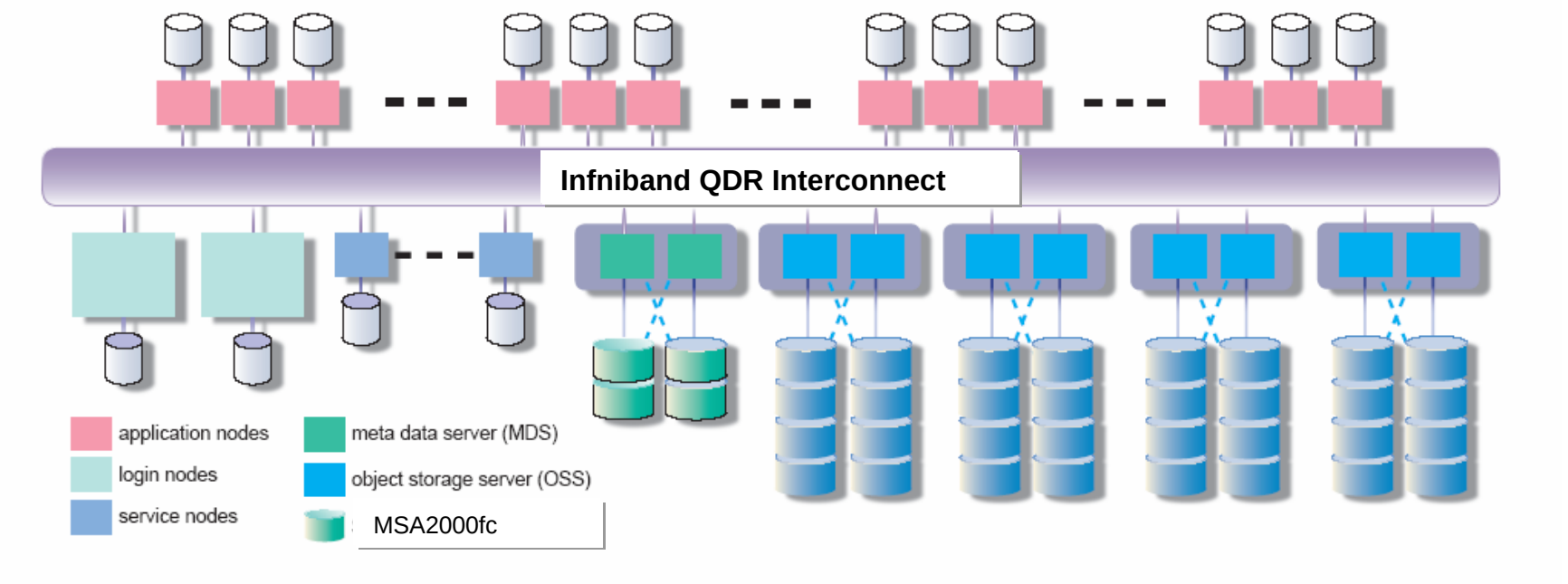
- Being commissioned in June 2010
- 448 Nehalem cores HP DL1000, 112 TeraByte online storage

Large Scale Data Facility: Digital Microscopy

- Combination of GPFS file system and Hadoop cluster
- Real-time data taking and analysis
- Implement data service for desktop users



HPC Cluster at KIT (HC3)

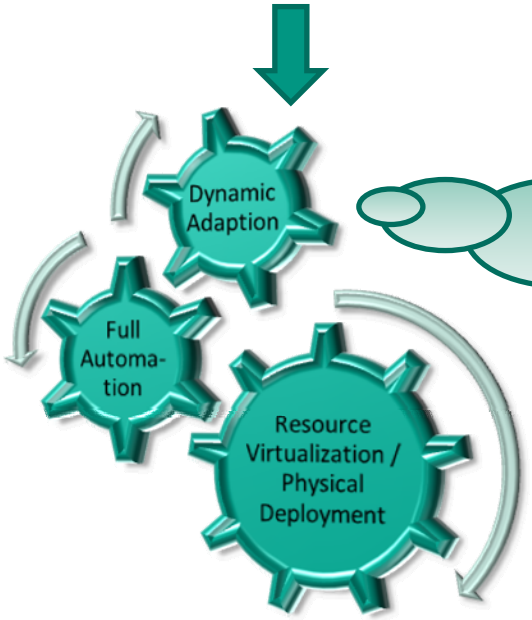


2656 Nehalem Cores / 332 Servers / 10 TB Memory
Infiniband QDR Interconnect
192 TB Lustre Filesystem

HPC as a Service

Traditional HPC Architecture …

- is characterized by very specific computer clusters designed for special applications
- offers pre-defined operation systems and user environments only
- serves one single application at a given time
- provides restricted user access
- provides management privileges exclusively to administrators

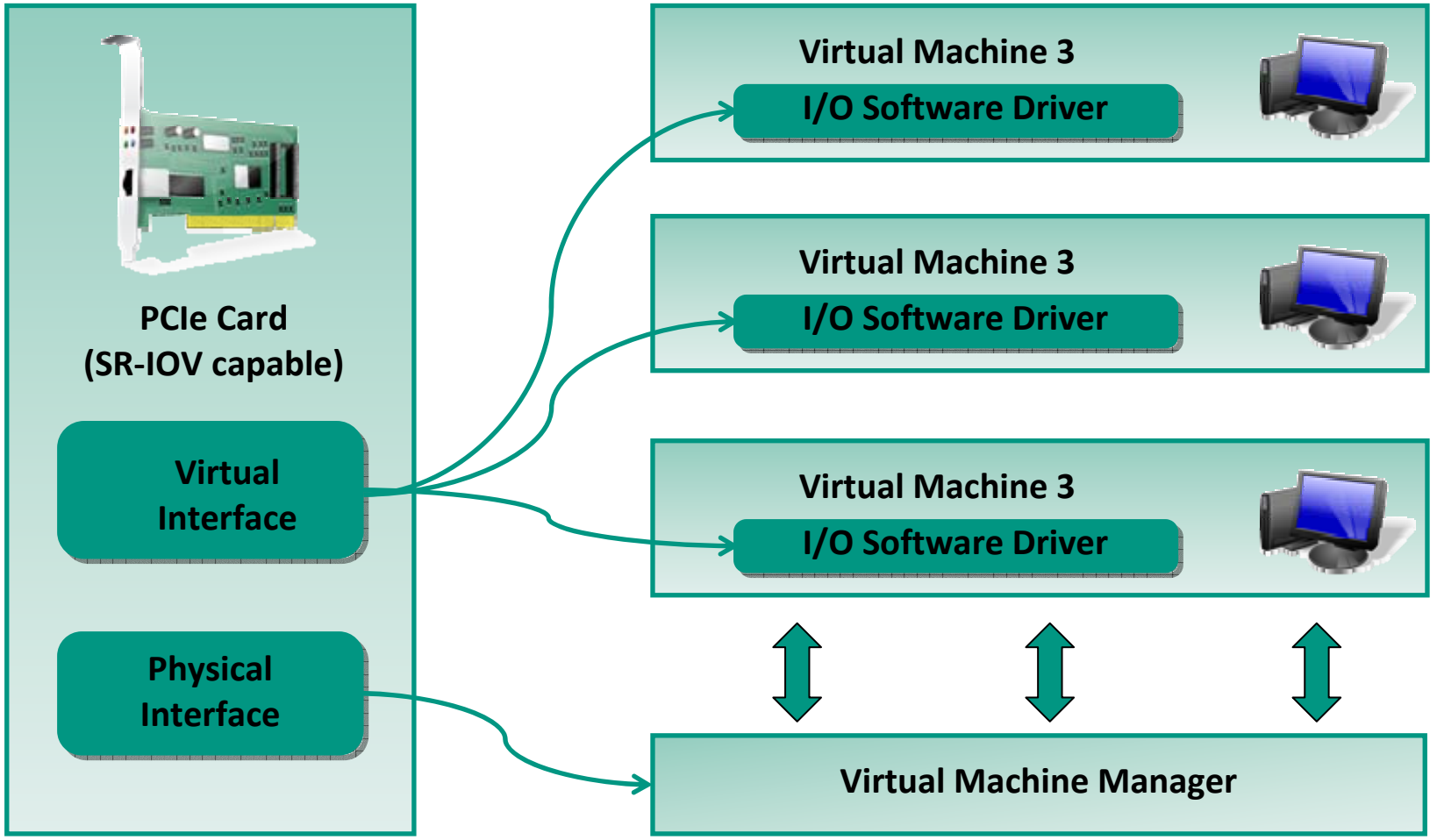


Concept of HPCaaS

- Capability of using clustered servers and storage as resource pools, fully automated management
- Individual cluster configuration on-demand
- Flexibility to serve multiple user groups and applications with varying requirements
- Customers gain resource management privileges

HPC as a Service: Virtualization of CPU and I/O

Infiniband Single Root – I/O Virtualization (SR-IOV)



InfiniBand Virtualization: SR-IOV

- **Limits of traditional I/O virtualization**
  - Increased I/O latency: VMM must process and route every data packet and interrupt, leads to higher application response time
  - Scalability limitations: software-based I/O processing consumes CPU cycles, reduces the processing capacity
- **Solution: SR-IOV provides near-native performance**
  - Extension to the PCI Express specification suite
  - Physical I/O resources are virtualized within the PCIe card, each card presents multiple virtual I/O interfaces
- **Virtual Functions (VF)**
  - Provide all the functionality which is necessary for communication
  - VMs interface directly with a VF without VMM intervention
- **Physical Functions (PF)**
  - VMM interfaces with PF to configure and manage I/O resource sharing among the multiple VMs

Current InfiniBand Development

- First SR-IOV supported IB Host Channel Adapters (HCAs) are already available by Mellanox® Technologies:
  Model Type: **ConnectX®-2**

- SR-IOV supported Drivers for the OFED Software Stack and Firmware are currently in development and will be available in the second half of 2010

Outlook

- The SR-IOV standard will allow to use InfiniBand in virtual machines at near-native performance
- Cloud frameworks like Eucalyptus or OpenNebula have to be extended to work with virtual resource bundles
- **Customers will be able to configure and deploy complete HPC systems on-demand according to their specific requirements!**

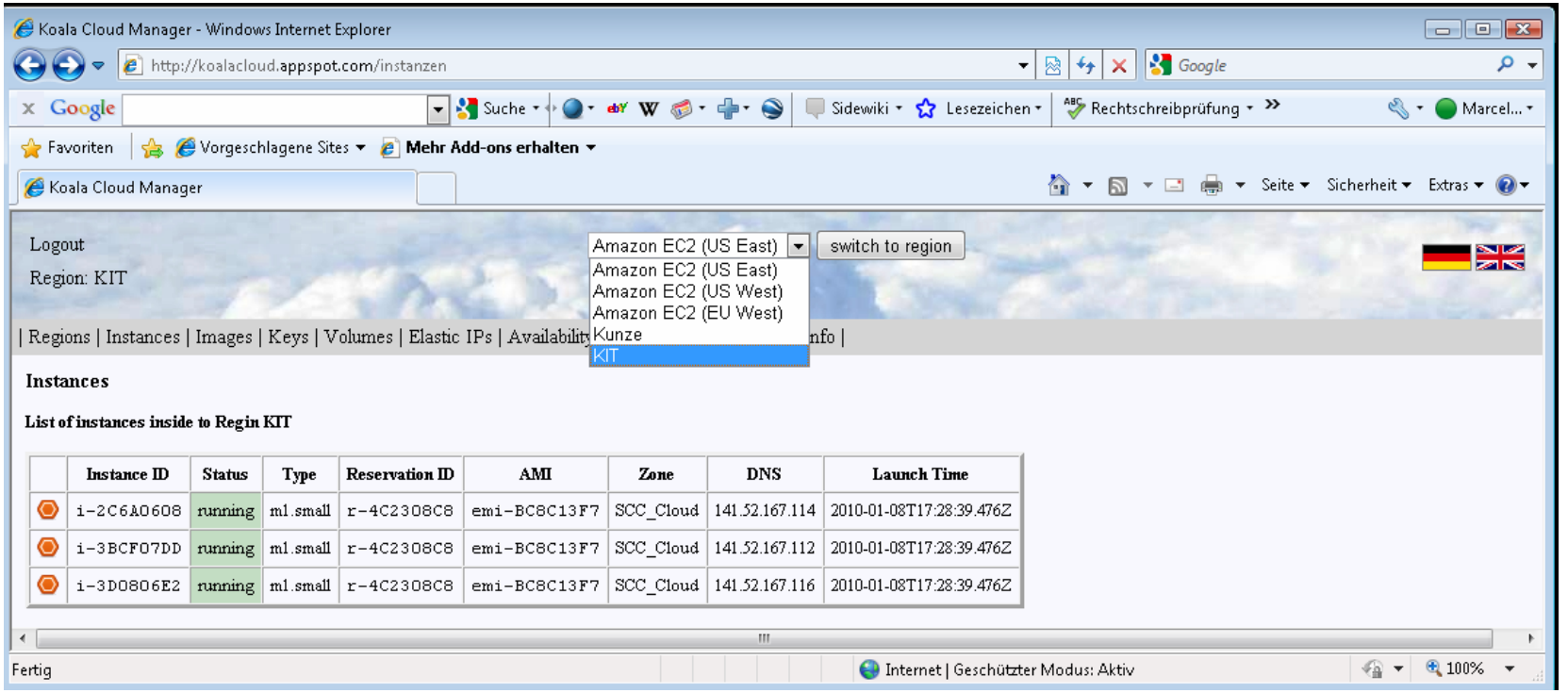
Cloud Management Tools

- **Three categories of management tools exist**
  - Browser based (e.g.: Elasticfox, Hybridfox)
  - Proprietary SaaS by provider (Amazon AWS console)
  - Proprietary command line based

- **The optimum tool would**
  - Run out of the cloud as a service (No local installation)
  - Run with any web browser (Pure HTML)
  - Support standard protocols like EC2/S3/EBS and OCCI
  - Be Open Source to cover a broad range of products

  - **KOALA: Karlsruhe Open Application (for) cLoudAdministration**

Cloud Management with KOALA

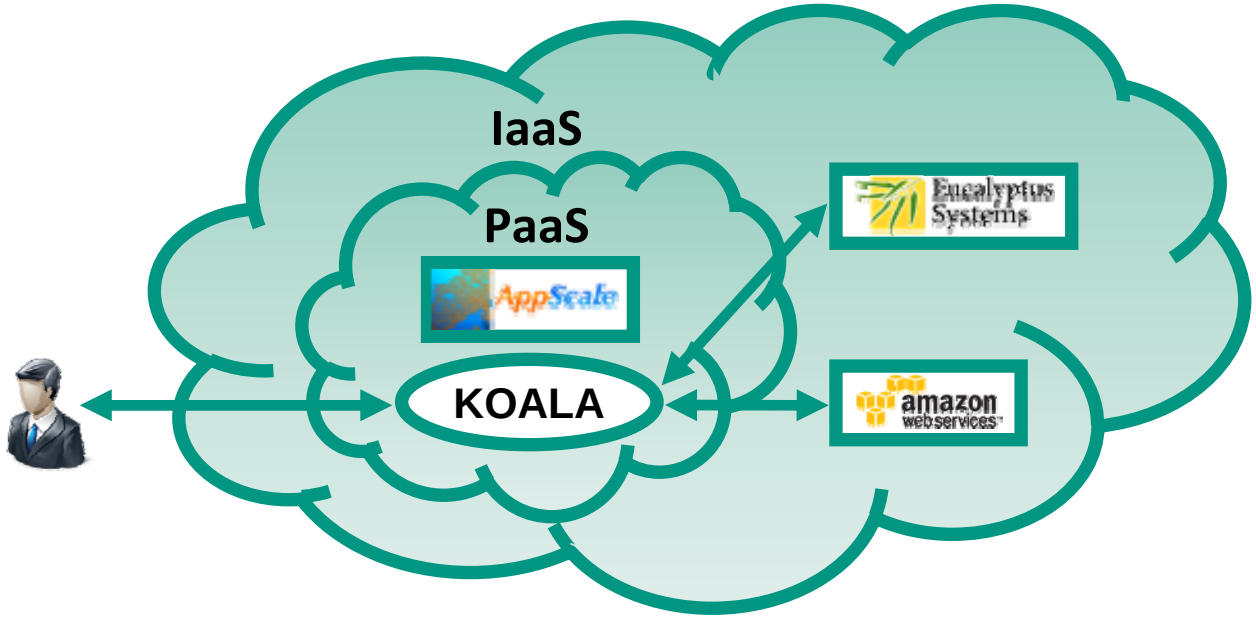


- **Management of public and private cloud services in a hybrid cloud: AWS, Eucalyptus, Nimbus, Tashi, ONE, …**
  - Runs as a software service in Google App Engine
  - Multi tenant management of CPU, storage and network
  - Future perspective: Dynamic management of resource bundles
    - MPI and Hadoop clusters in the cloud

KOALA running inside a Private Cloud PaaS

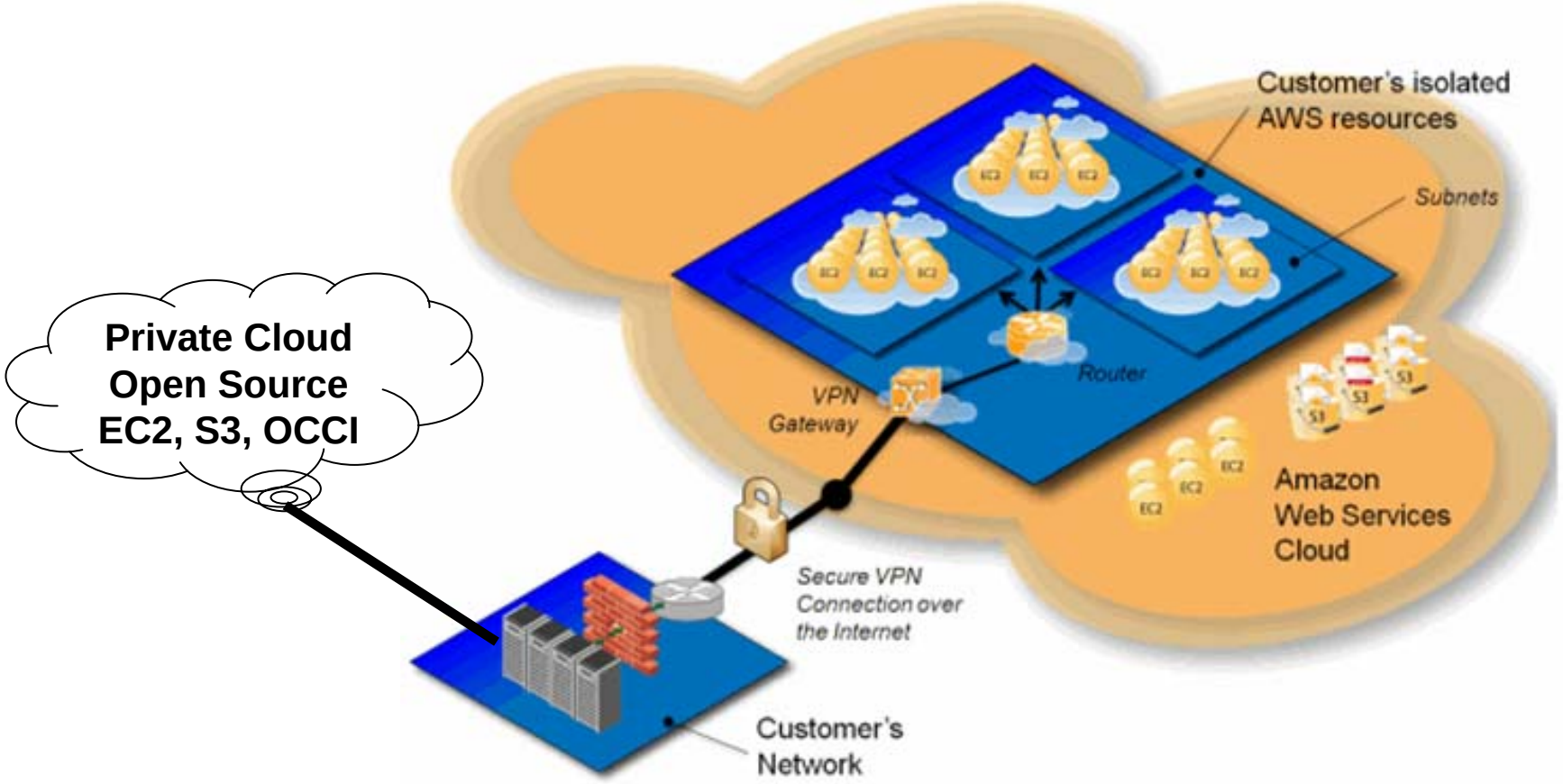
http://koalacloud.appspot.com - http://code.google.com/p/koalacloud

- **KOALA also runs in a Private Cloud PaaS based on AppScale**
  - AppScale is an Open Source re-implementation of the Google App Engine



- **Allows users to do cloud management inside their own Private Cloud in case ofconcerns regarding privacy and availability**

Hybrid Clouds: Cloud Bursting



- Transfer workloads and data transparently between clouds
- Example: Amazon Virtual Private Cloud

Summary

- **Open Cirrus project offers interesting R&D opportunities**
  - Cloud systems development
  - Cloud application development
  - Accepting research proposals
- **Open Source cloud stack**
  - Standards are important to foster a cloud computing market
  - De-facto standards are around (e.g. Amazon EC2, S3)
  - Standards are evolving (OCCI)
  - Federation of private and public clouds seems important (Cloud bursting)


©johnlund.com