CloudStack技术沙龙

北京站 2012/10/18

内容安排

- 总体概述
- CloudStack Business in China
- Scalability
- 深入管理服务器
- CloudStack4.0新增功能介绍
- CloudStack vs. OpenStack

cloudstack

总体概述

- CloudStack是什么
- CloudStack的特点
- 总体架构

What is CloudStack?

- Secure, multi-tenant cloud orchestration platform
  - Turnkey platform for delivering IaaS clouds
  - Hypervisor agnostic
  - Scalable, secure and open
  - Open source, open standards
  - Deploys on premise or as a hosted solution
- Deliver cloud services faster and cheaper

CloudStack是什么

云环境中低成本的对资源综合管控的平台

- 云
  - 公有云
  - 私有云
  - 混合云
- 资源
  - 计算
  - 存储
  - 网络

三种云

公有云

- 多租户
- 共享资源
- 弹性扩展
- 按需付费
- 公共网络

混合云

- 企业托管
- 专有资源
- 安全性
- SLA
- 第三方运维

私有云

- 企业内部
- 专用资源
- 安全 性
- 完全控制
- 内部网络
- 企业直接管理

CloudStack的特点

- 异构Hypervisor
- 异构存储
- 丰富的网络功能
- 强大的扩展能力
- 与实体环境一致的设计思想
- 安全性
- 功能架构

异构及网络

计算

Hypervisor

XenServer | VMware | Oracle VM | KVM | Bare metal

存储

块 & 对象

Local Disk | iSCSI | Fiber Channel | NFS | Swift

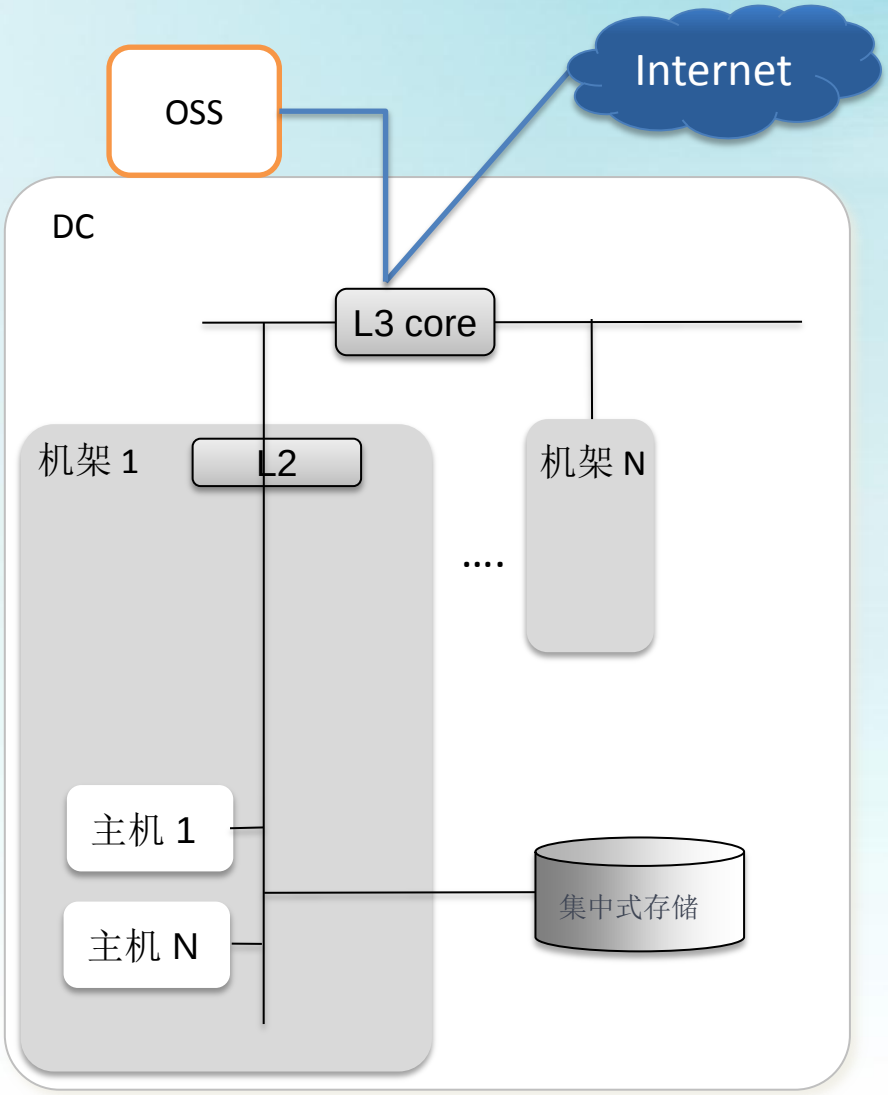
主存储 | 二级存储

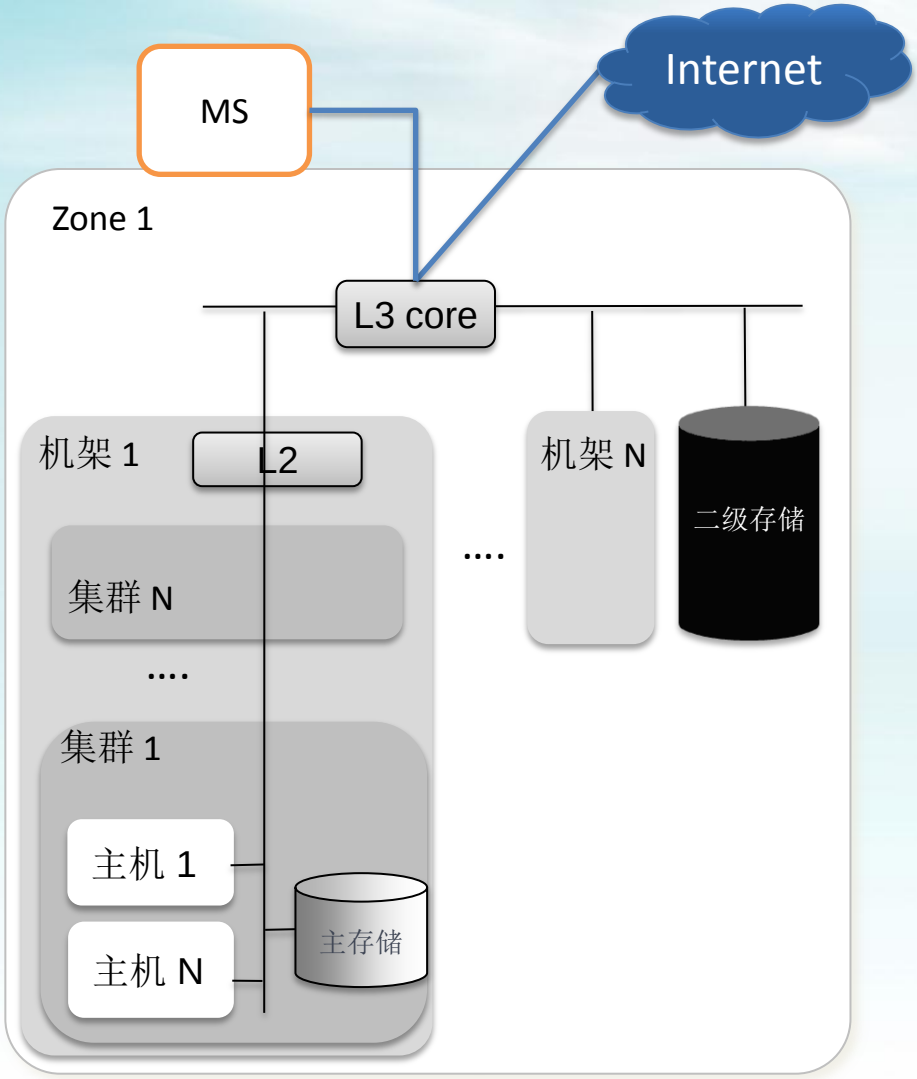
网络

网络 & 服务

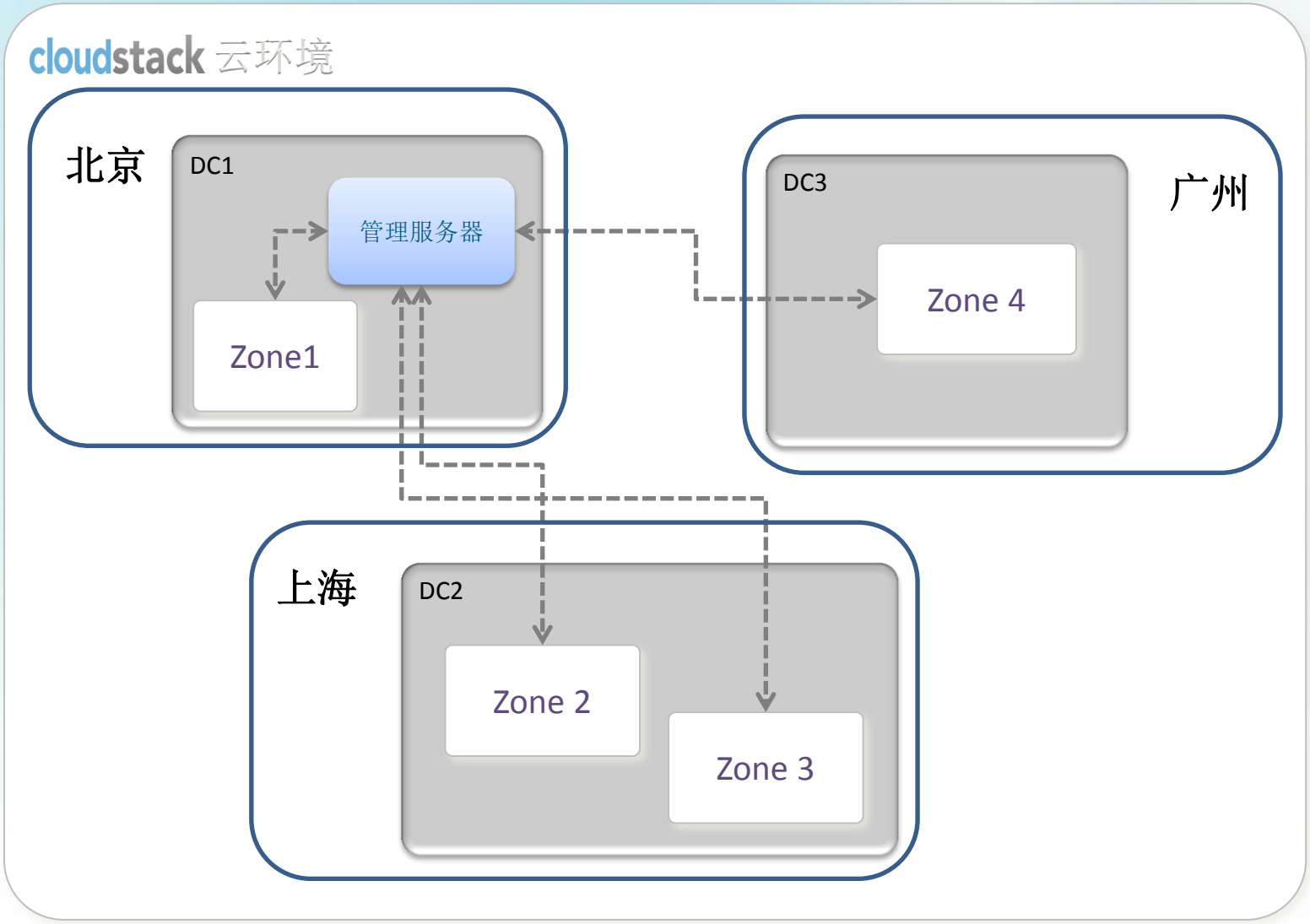
TC | VLAN | Firewall | LB | VPN | S/DNAT | ...

扩展性及设计


OSS
Internet
DC
L3 core
机架 1
L2
机架 N
....
主机 1
主机 N
集中式存储


MS
Internet
Zone 1
L3 core
机架 1
L2
机架 N
二级存储
....
集群 N
....
集群 1
主机 1
主机 N
主存储

扩展性及设计(续)


cloudstack 云环境
北京
DC1
管理服务器
Zone1
DC3
广州
Zone 4
上海
DC2
Zone 2
Zone 3

安全性

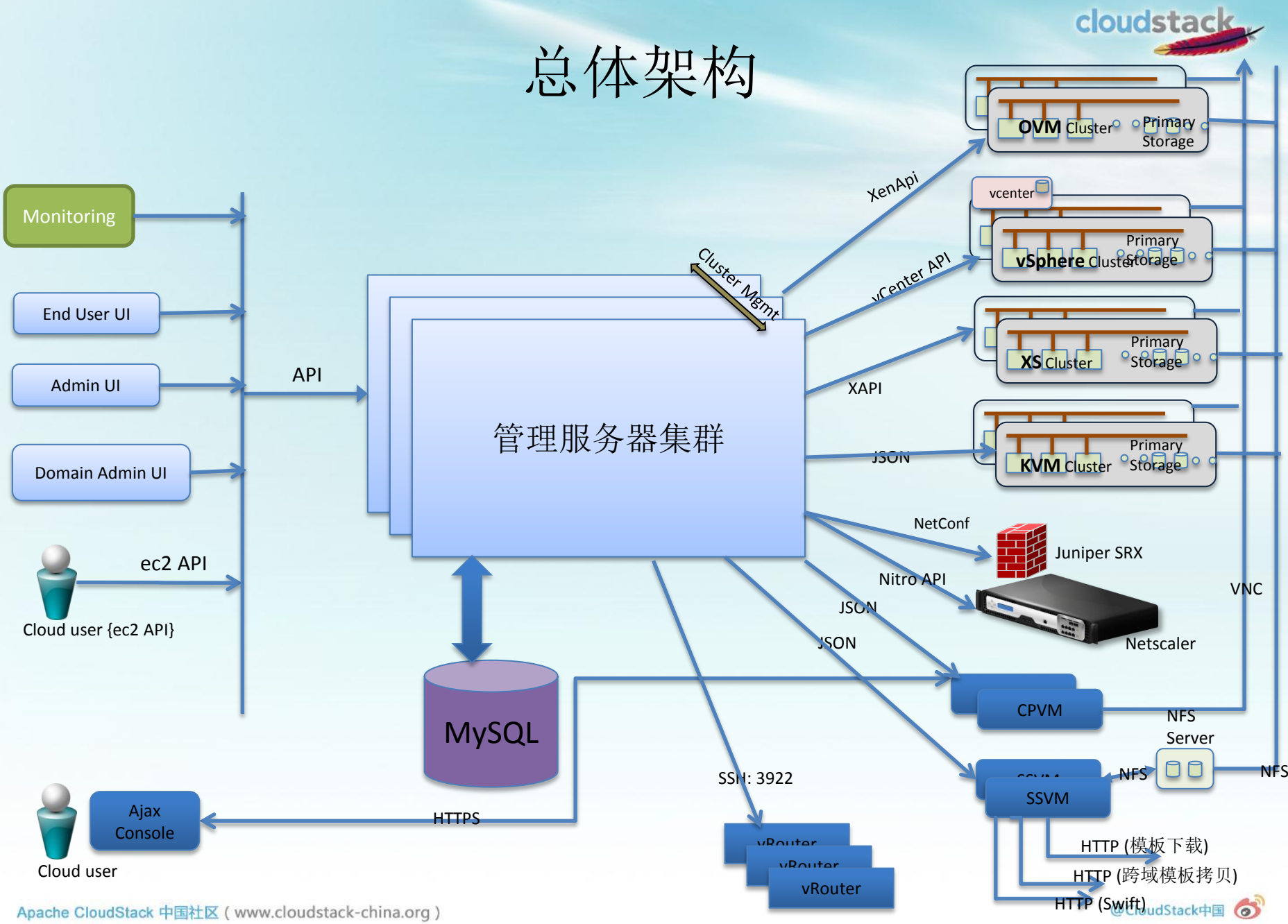
CloudStack通过一系列密码/密钥来提供安全防护或权限识别,这些密码/密钥被自动加密保存:

/usr/bin/cloud-setup-databases

/etc/cloud/management/db.properties

- Database secret key
- Database password
- SSH keys
- Compute node root password
- VPN password
- User API secret key
- VNC password

总体架构


cloudstack
Monitoring
End User UI
Admin UI
Domain Admin UI
API
ec2 API
Cloud user {ec2 API}
管理服务器集群
Cluster Mgmt
MySQL
Ajax Console
Cloud user
HTTPS
XenApi
vCenter API
XAPI
JSON
OVM Cluster
Primary Storage
vcenter
vSphere Cluster
XS Cluster
KVM Cluster
NetConf
Juniper SRX
Nitro API
JSON
JSON
Netscaler
VNC
CPVM
NFS Server
NFS
NFS
SSVM
SSH: 3922
vRouter
vRouter
vRouter
HTTP (模板下载)
HTTP (跨域模板拷贝)
HTTP (Swift)

CloudStack Business in China

CloudStack Scalability

cloudstack

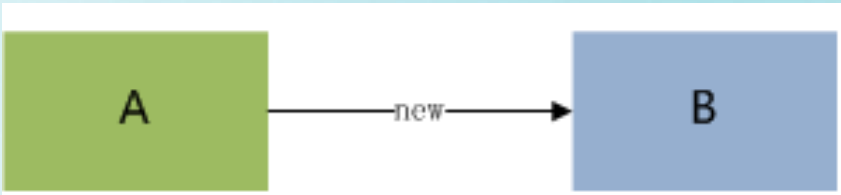
管理服务器

- 内部原理
- 分层结构
- 启动及处理流程
- Adapter
- Interceptor
- DAO
- 类命名规则

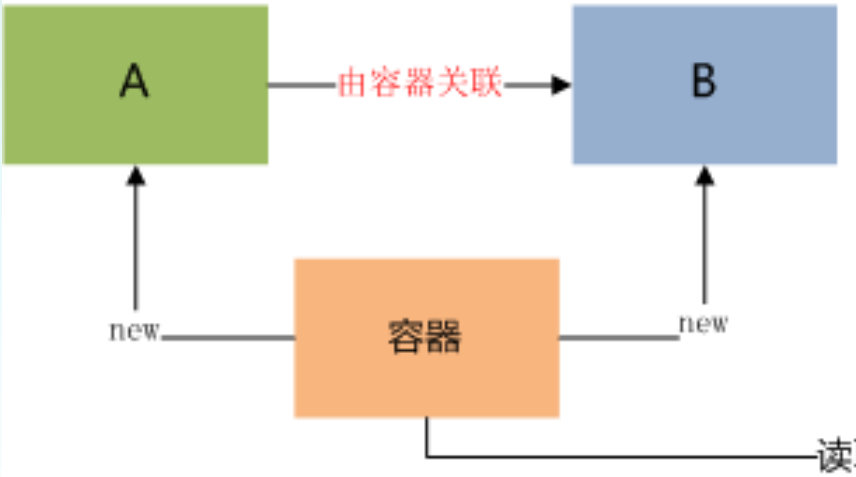
内部原理

- Tomcat webapp
- Java 单进程
- IoC, 实现类似Spring功能
- cglib实现AOP
- ORM, 实现类似Hibernate功能

IoC (依赖注入)




```
Class A {
        private B b = new Bimpl();

}
```




```
<adapter name="a" class="com.tcloud.A"/>
<adapter name="b" class="com.tcloud.BImpl"/>

Class A {
        @inject
        private B b;

}
```

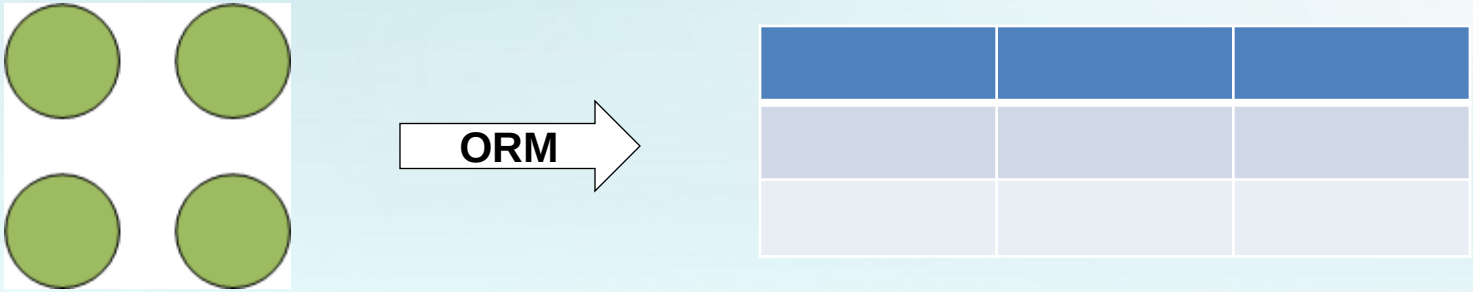
AOP (面向方面编程)


```
public void doSomething(){
        logger.debug(...);
        authentication.check(..)
        txn.start(...);

        Realfunction()

        txn.close(...);
}
```


```
public void doSomething(){
        Realfunction()
}

Method interceptor
XML definition
```

ORM (关系对象映射)

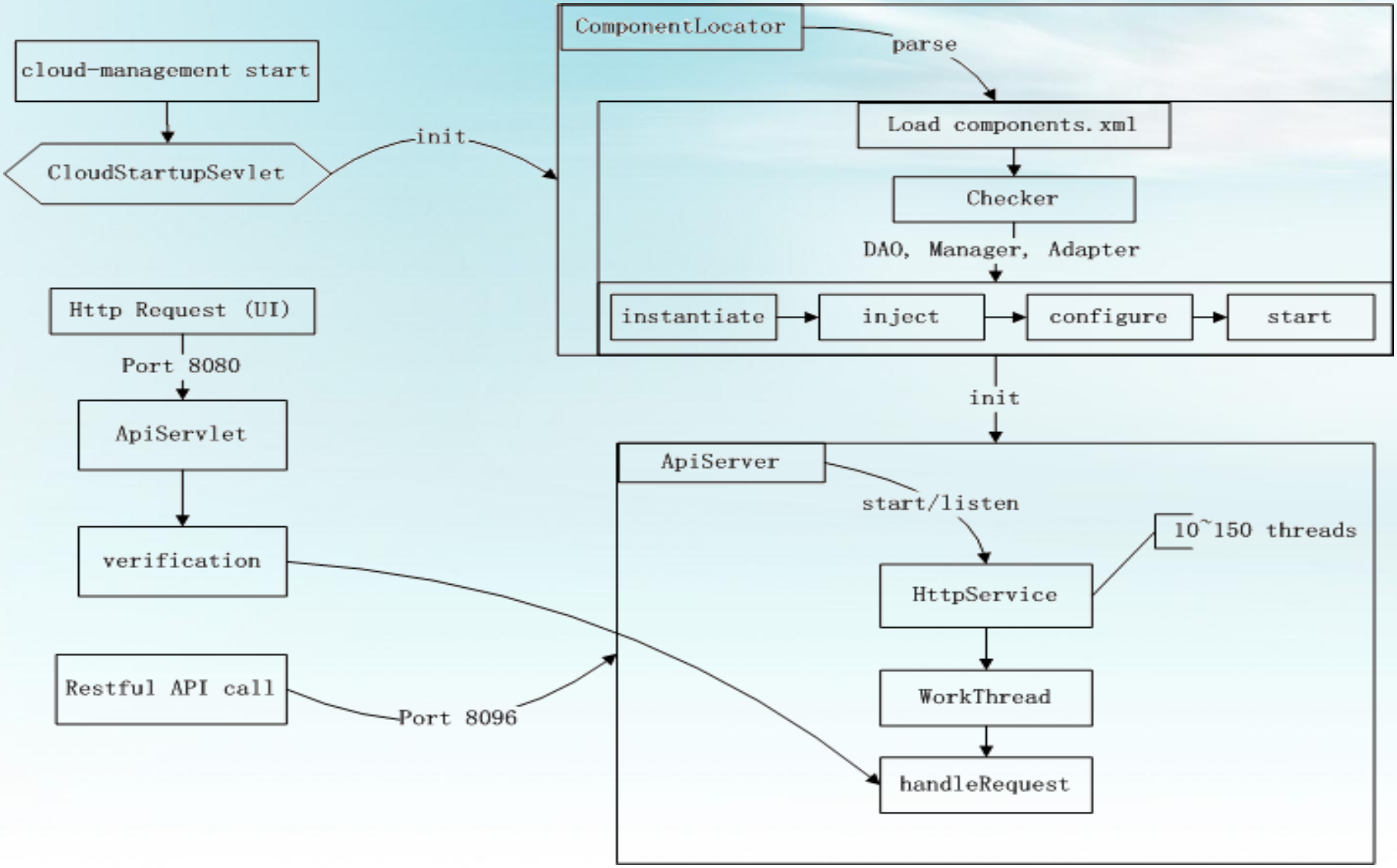
- 数据库为关系型，和对象不能直接映射
- ORM提供根据对象对关系数据库进行增删查改操作的功能



分层结构

Services	• 实现了所有CloudStack HTTP API
Management	• 封装业务逻辑 • 与Adapter交互实现定制化功能
Resources	• 封装了CloudStack对各类物理资源（和hypervisor）的调用
Data Access	• 封装数据访问层

启动及处理流程

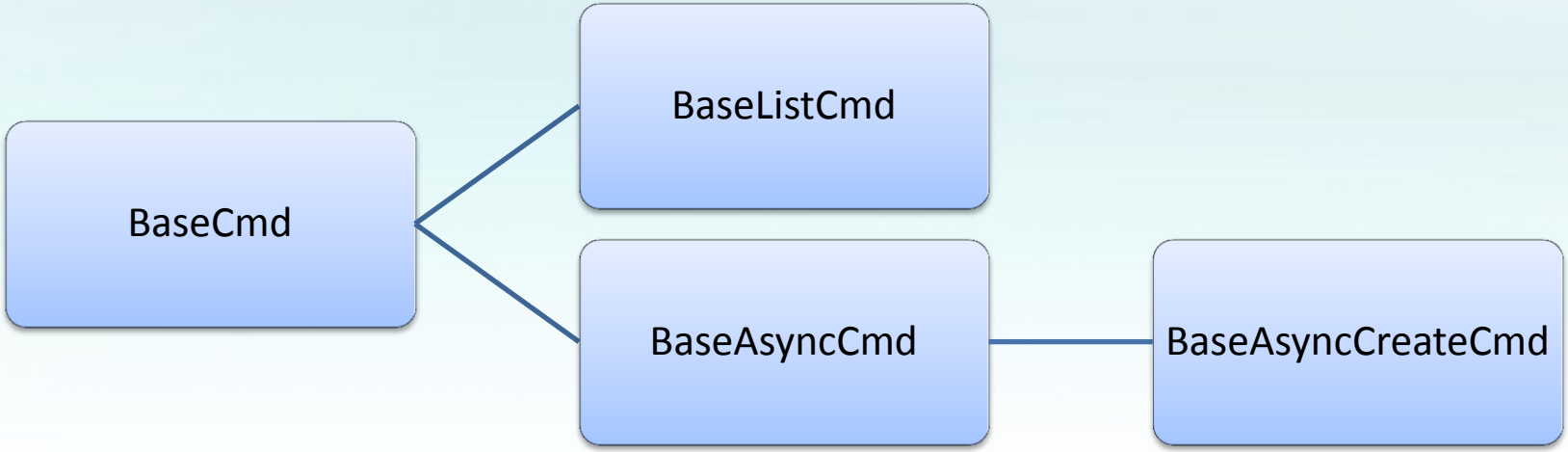

cloud-management start
CloudStartupSevlet
init
ComponentLocator
parse
Load components.xml
Checker
DAO, Manager, Adapter
instantiate
inject
configure
start
Http Request (UI)
Port 8080
ApiServlet
verification
Restful API call
Port 8096
init
ApiServer
start/listen
10~150 threads
HttpService
WorkThread
handleRequest

Cmd

/etc/cloud/management/commands.properties

listAccounts=com.cloud.api.commands.ListAccountsCmd;15

1	2	4	8
ADMIN	RESOURCE_DOMAIN_ADMIN	DOMAIN_ADMIN	USER

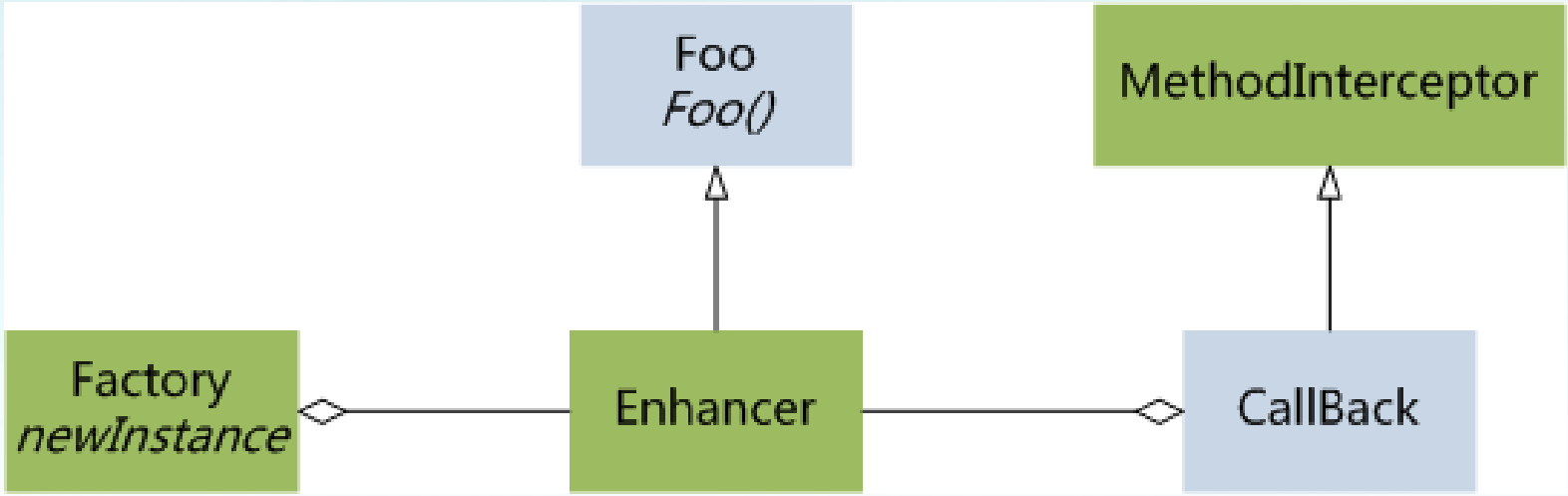


Adapter

- 目标：模块化、可扩展、可替换
- 分类
  - Discoverer
    - XcpServerDiscoverer, SecondaryStorageDiscoverer,…
  - Allocator
    - RecreateHostAllocator, FirstFitAllocator,…
  - Investigator
    - VmwareInvestigator, XenServerInvestigator,…
  - XXXXGuru
    - HypervisorGuru, NetworkGuru
  - Listener
    - StorageSyncListener,LocalStoragePoolListener,…
  - …

Intercepter

- Proxy模式
- 使用cglib库
- ActionEventCallback
- DatabaseCallback



DAO

- GenericDaoBase
  - 定义common的方法,CRUD
  - 拦截器UpdateBuilder用于记录VO对象的字段变更
  - 拦截器SearchBuilder用于初始化search语句查询条件
  - SearchCriteria用于指定查询参数
  - Filter用于order by和limit

DAO

- 如何写一个Dao
  - @Local
  - extends GenericDaoBase<K,V>
  - 构造函数中定义SearchBuilder，运用and, or, join,groupby以及操作符初始化
  - 实现方法，在方法中使用SearchCriteria传入参数
  - 通过使用Filter进行排序和limit
- VO操作
  - createForUpdate
  - New

DAO 实例

```
// ExampleVO.java
@Entity
@Table(name="example")
public class ExampleVO {
    @Id
    @GeneratedValue(strategy=
GenerationType.IDENTITY)
    @Column(name="id")
    long id;

    @Column(name="name")
    String name;

    @Column(name="value")
    String value;
}
```

```
//  ExampleDao.java
public interface ExampleDao
    extends GenericDao<ExampleVO,
Long> {
}

//  ExampleDaoImpl.java
@Local(value=ExampleDao.class)
public class ExampleDaoImpl
    extends
GenericDaoBase<ExampleVO, Long>
    implements ExampleDao {

    protected ExampleDaoImpl() {
    }
}
```

类命名规则

- 接口实现的后缀名Impl
- 对数据库表映射的类，后缀VO
- 为API层提供服务的类，后缀Service
- 数据访问类，后缀Dao
- API对应的实现，后缀Cmd
- 业务逻辑和resource之间的通讯，后缀Command
- 物理资源实现，后缀Resource

cloudstack

4.0新增功能

- VLAN之间路由 (VPC)
- Site-to-Site VPN
- 数据盘支持本地存储
- 虚拟资源Tagging
- HA专用主机
- 支持 AWS API
- 支持 Nicira NVP (L2)
- 支持Caringo作二级存储
- KVM Hypervisor 支持升级到 Ubuntu 12.04 和 RHEL 6.3

VPC (inter-VLAN Routing)

- 网络ACL允许的情况下, 在1-N的VPC网络里部署虚机, 虚机里属于N-1的网络可以和属于N的网络通过虚拟路由器进行通信
- 账户A创建的VPC, 只允许账户A的网络加入到这个VPC
- 管理员为账户A创建的VPC, 只允许账户A的网络加入到这个VPC
- 管理员/用户可以添加静态路由(CIDR + GW)来控制网络流量到下列的目的地:
  - —VPN 网关
  - —私有网关
- 网络ACLs - 通过添加网络规则来控制允许/拒绝来宾网络之间的流量

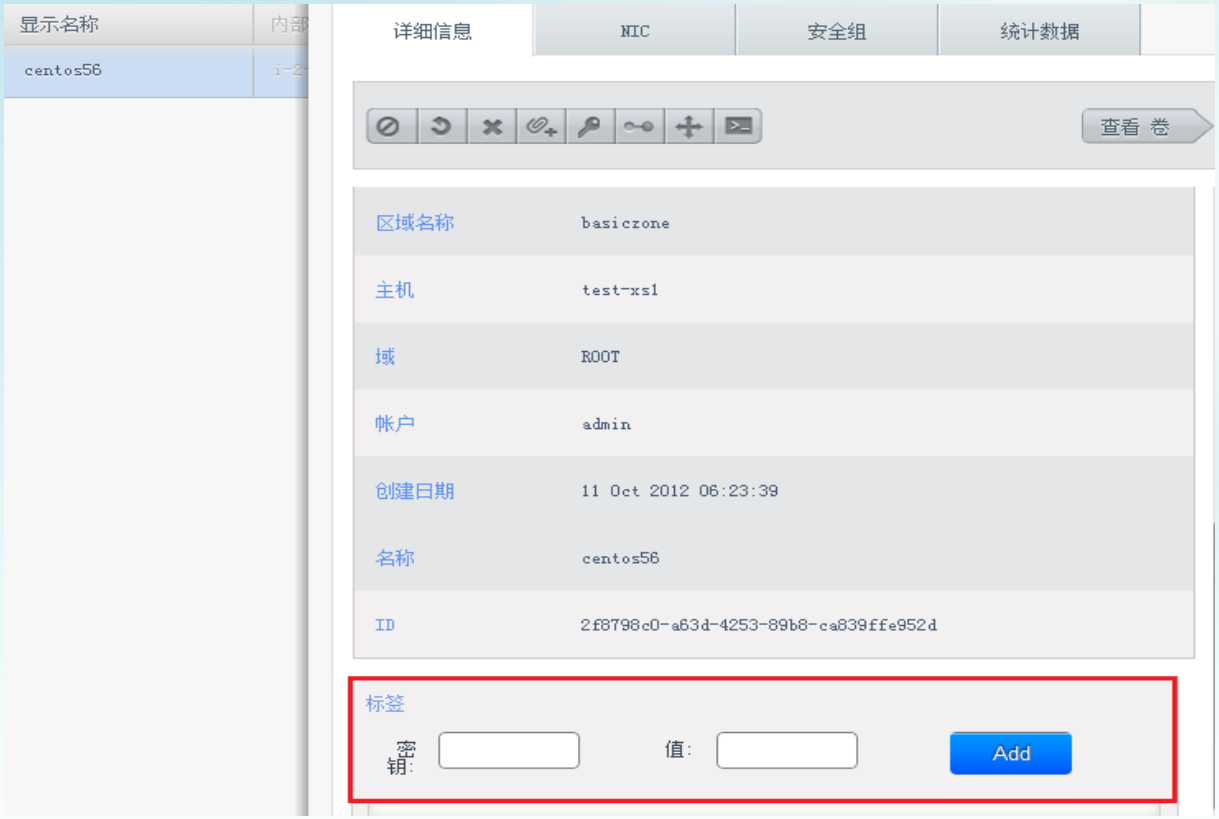
VPC—配置


控制板
实例
存储
网络
模板
事件
帐户
域
基础架构
项目
全局设置
服务方案
网络 - VPC
配置 VPC: MyVPC
MyVPC
虚拟路由器
Applicat...
0 VM
192.168.3.0/24
DB Tier
3 VM
192.168.2.0/24
Web Tier
1 VM
192.168.1.0/24

Tagging

允许用户保存各种资源的元数据信息, AWS风格, Key-Value对, 凡是有Object ID的资源都可以设置

- User Vm
- Template
- ISO
- Volume
- Snapshot
- Guest Network
- LB rule
- PF rule
- Firewall rule
- Security Group
- Public IP Address
- Project
- Vpc
- NetworkACL
- StaticRoute



Caringo Castor

Caringo提供的云产品: Elastic Content Protection, CloudScaler 和Indexer; 基于Castor.

Caringo 自身保证高性能, 简单管理以及云模块之间的互操作性, 同时降低云环境中的复杂性. Caringo可以在任意的存储硬件上部署

Nicira NVP

- 每个租户任意数量的虚拟私有网络—没有VLan
- 持续的高可用性,包括分布式active-active的集群故障转移
- 网络硬件无关
- 数据中心互连互通允许虚拟网络扩展到多个数据中心或资源域
- 整个云环境中网络服务的可编程及自动化
- 为大规模,生产环境云平台提供运维工具

Nicira NVP - TODO

Nicira目前没有GUI可用,目前支持的API:

- addNetworkServiceProvider
- updateNetworkServiceProvider
- addNiciraNvpDevice

第二阶段工作:

- 集成L3 support
- 集成NetworkOfferings

cloudstack

cloudstack vs openstack