

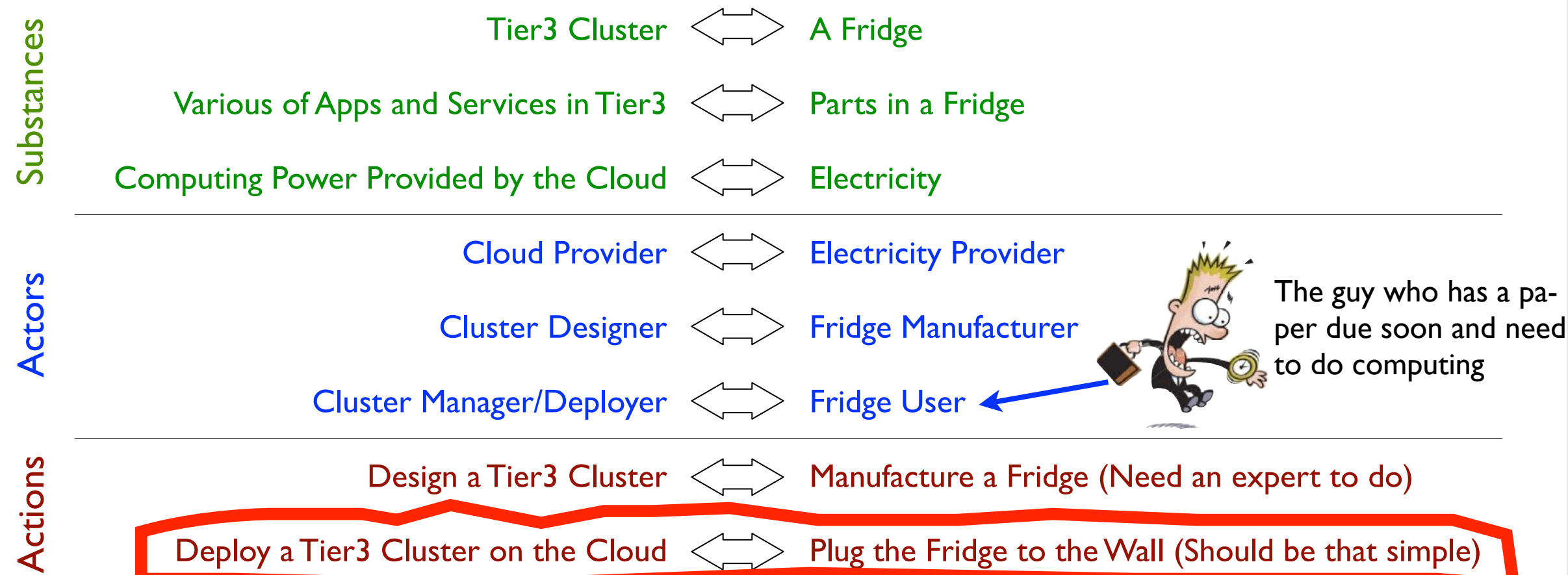
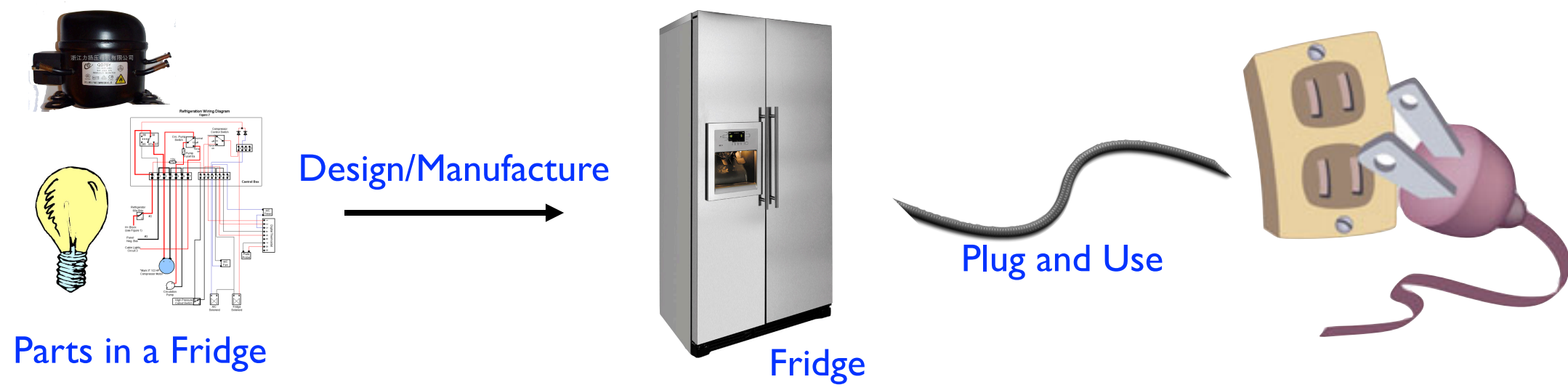
CloudCRV and Virtual Cluster Appliance

Ship Your Cluster to the Cloud with 1-Click



CloudCRV, the Idea

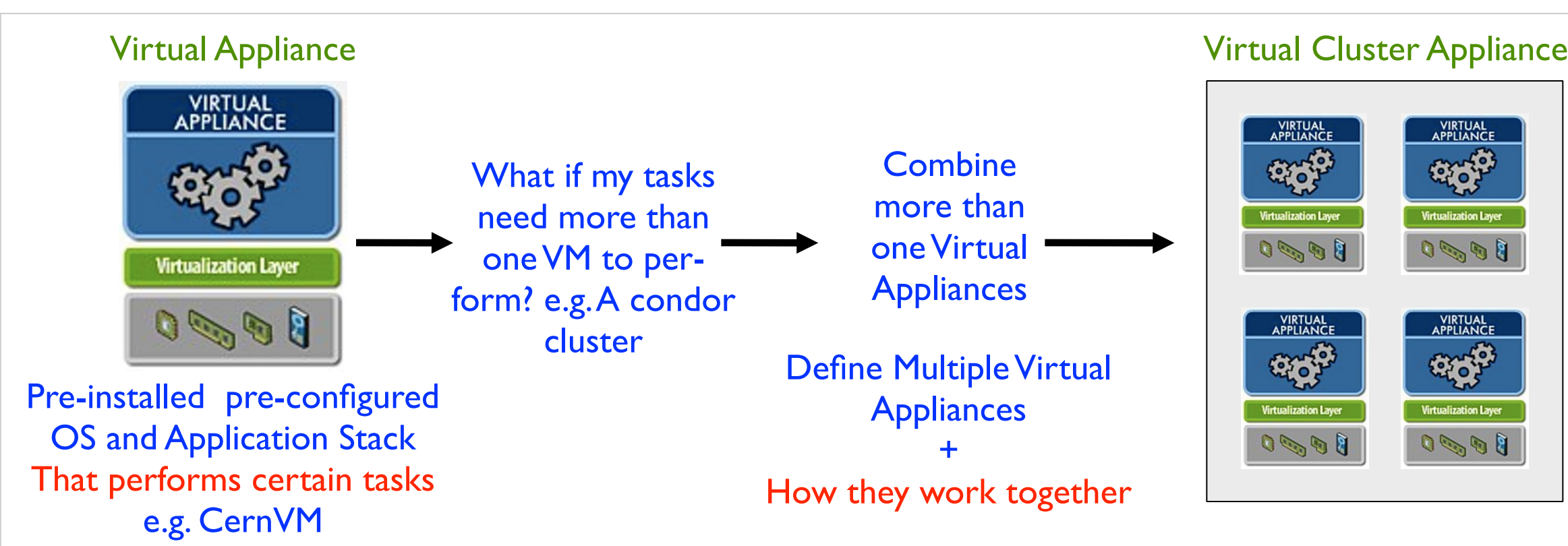
The Utility-Cloud Analogy



◆ **Ideally**: if you have a conference paper due next Monday, you can allocate many machine to redo your calculation in 1 day (comparing to your jobs waiting in the batch queue of your institution for weeks)

Yes, it should be that Simple

From Virtual Appliance to Virtual Cluster Appliance

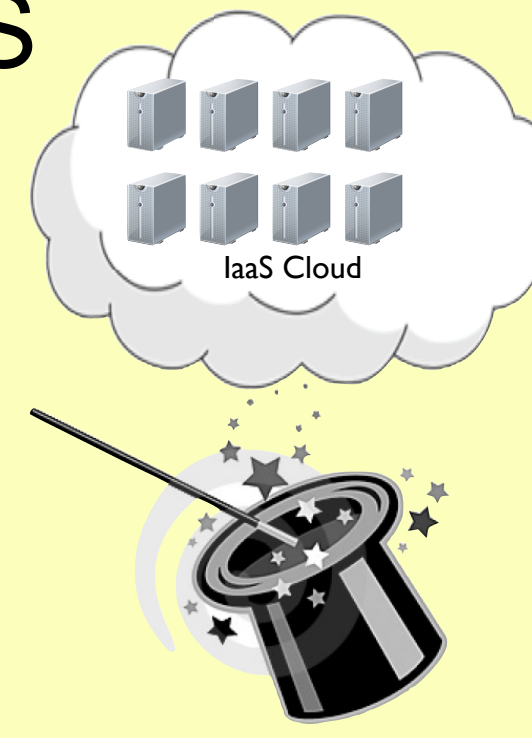


Virtual Cluster Appliance = Multiple Virtual Appliances + Their Relationship + Deploy Time Instruction

- Design**
 - ◆ Virtual Cluster Appliances are Designed by Experts:
 - ◆ Install and Configure all the software and services
 - ◆ Define how the different VMs and services interact during deployment so that they can work together.
- Distribute**
 - ◆ Virtual Cluster Appliances are Packaged into a set of Scripts and specifications in an SVN Repository
 - ◆ Which base image to start from How to install/configure services to work together
 - ◆ We use **Puppet** to help the building/configuration process
- Deploy**
 - ◆ Virtual Cluster Appliances are Deployed by non-Experts:
 - ◆ Start CloudCRV and point to the VCA's SVN Repository
 - ◆ Give basic setup parameters, e.g. Amazon account info, how many worker nodes are needed

The Useful Cloud

◆ With the help of IaaS cloud (Such as Amazon EC2), anyone can lease a large amount of Computing Resource (e.g. 100 Linux Computers)



Really Useful?

Right, I can get many linux computers now, but **how do I configure them** to do what I want?

- ◆ E.g. Tier3 Cluster with all ATLAS Software, Job Scheduling, Obtain Data automatically?
- ◆ I don't know how to set them up
- ◆ Even if I do, do it by hand is too much work

Problem:

There is no tool to automatically deploy a cluster onto the cloud and configure all the VMs to work together.

Solution - CloudCRV

- ◆ The experts design a Virtual Cluster Appliance (VCA)
- ◆ With the help of CloudCRV, I can deploy the VCA to the cloud with minimal customization + 1-click
- ◆ CloudCRV can **scale** the size of the cluster on demand

Want to Try it? Contact Us

Questions or Comments to:

Yushu Yao, yao.yushu@gmail.com

Website:

◆ <http://code.google.com/p/cloudcrv/>

◆ First Entry if you Google it.

Try it:

◆ For new adopters, we can help you to setup a VCA that will fit your application. (We can easily do that since VCAs are highly reusable)



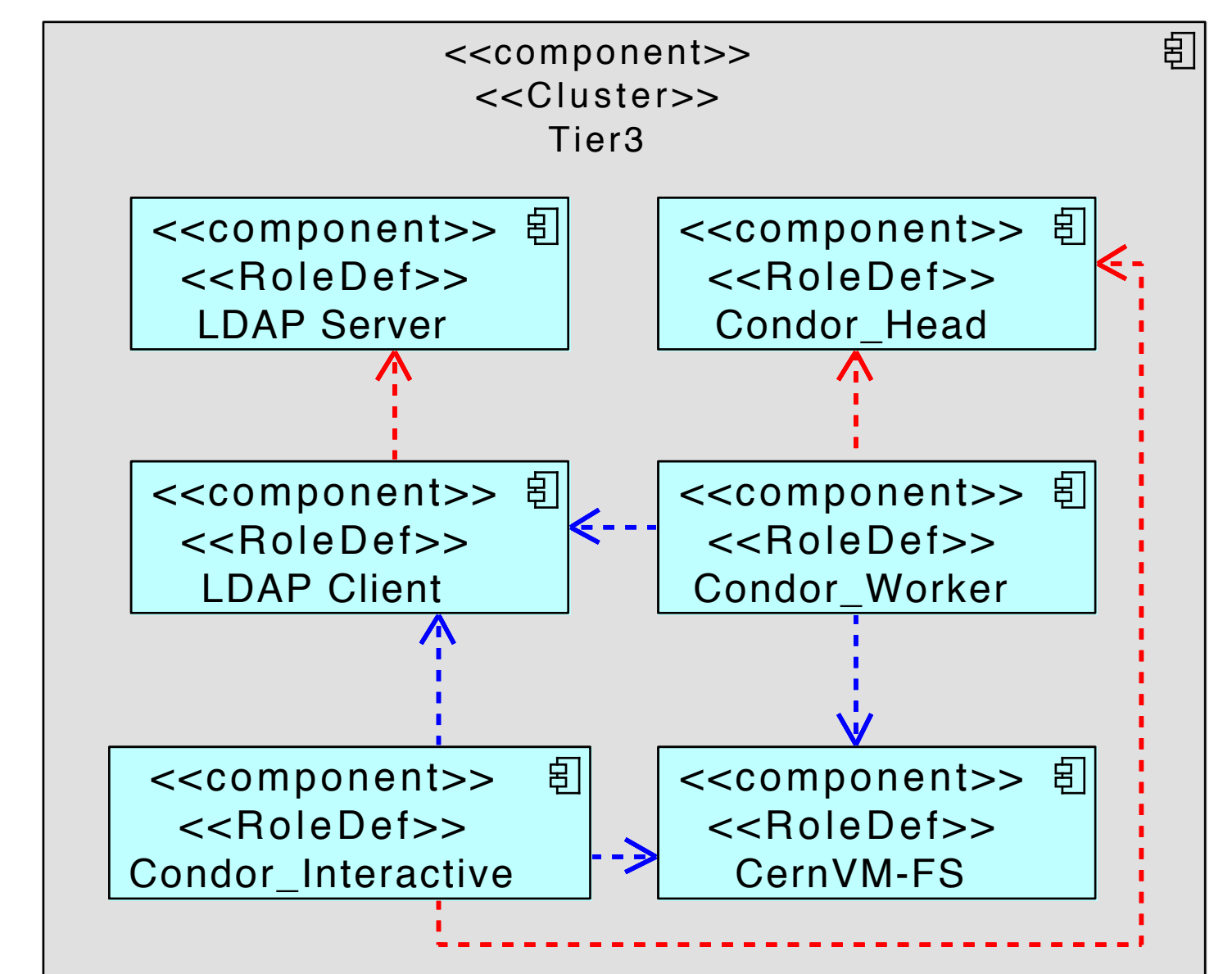
ATLAS Tier3, An Example

Components in CloudCRV

CloudCRV = Cloud + Cluster + Role + VM

- ◆ **Role**: One Basic Service (One Part in the Fridge)
- ◆ **Cluster**: A Collection of Roles
- ◆ **VM**: The Virtual Machine where the Roles are realized

Roles in a Minimal ATLAS Tier3

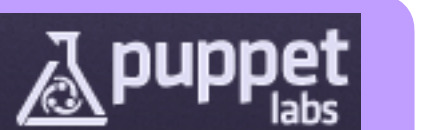


Shown above are the Roles of a simplified version of an ATLAS Tier3 Cluster:

- ◆ **LDAP Srv/Client**: Provides Authentication
- ◆ **Condor**: Job Submission/Execution
- ◆ **CernVM-FS**: Provide ATLAS Software

Arrows show the dependencies between Roles
Red=Remote: Two Roles are in different VMs
Blue=Local: Both Roles sit on the same VM

Defining Roles With Puppet

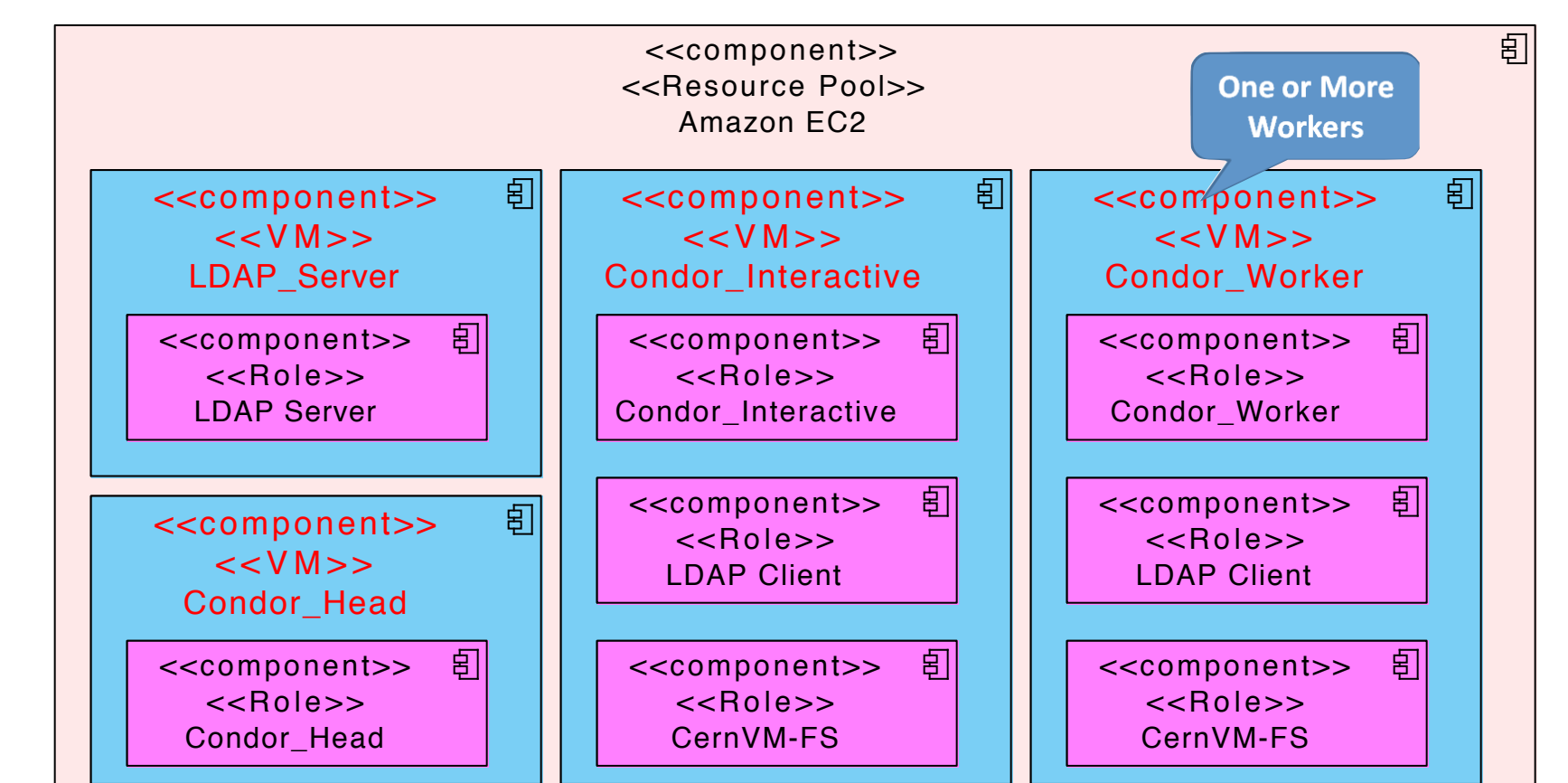


Roles are defined with Puppet - a tool that:
◆ Defines the target state of the machine with a high level language
◆ After all Roles are defined, we put them into an SVN repository. E.g. <https://atlastier3cfg.googlecode.com>

At Deployment Stage

Job of the Cluster Managers (Non-Experts):
◆ Start up CloudCRV server, which is normally a Linux VM and works out of the box
◆ Point to the SVN Repository
◆ Specify his Amazon EC2 account number and how many Condor Worker nodes he would like to use
◆ Click Start Button and Use the Cluster

The Deployed Cluster



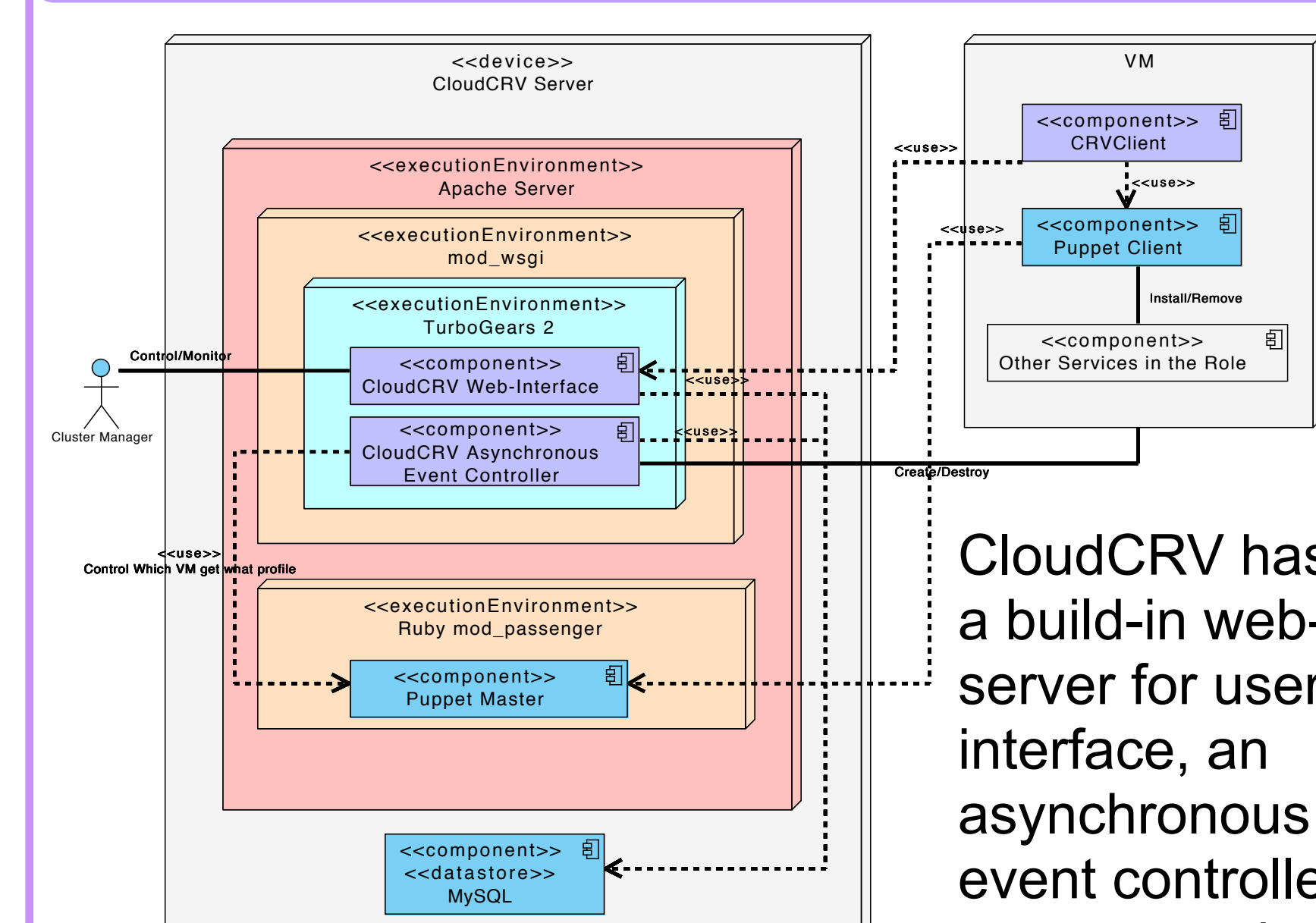
CloudCRV will start the VMs according to the definitions in the VCA. Only takes several minutes.

Auto Scaling

Pay-as-you-go and scaling On-demand are the greatest benefits we get from the Cloud.
◆ By talking with the batch scheduler (e.g. Condor Master), CloudCRV can fire up more Workers when the demand is high (e.g. when there are 100 jobs waiting but only 10 workers in the cluster)
◆ When demand is low, the Cluster size is reduced to save cost.

The Idea Want to know more? Example and Details

Inside CloudCRV



CloudCRV has a build-in web-server for user interface, an asynchronous event controller to manage the lifecycle of Clusters, Roles and VMs. A very thin client will run on each VM for communication and control purposes.