



Universidad Zaragoza

Diseño e implementación de un sistema dinámico de gestión de trabajos distribuidos en un entorno de máquinas virtuales

David Ceresuela

Proyecto fin de carrera – Ingeniería en Informática
Curso 2011/2012

Director: Javier Celaya

Introducción

- Aproximación clásica a la ejecución de trabajos:
 - *Cluster* de ordenadores
 - Computación en malla o *grid*
- Aproximación nueva a la ejecución de trabajos:
 - Computación en la nube
 - Acceder a una aplicación cuya lógica y datos están situados en una localización remota
 - Problema: administración de la infraestructura

Introducción

- Herramientas de gestión de configuración
 - Considerable avance en los últimos años
 - Describir y llevar a un sistema informático a un cierto estado
 - Iteración y convergencia
 - Capaces de administrar nodos en entornos heterogéneos y complejos
 - Incapaces de administrar infraestructuras distribuidas como una entidad propia

Puppet

- Lenguaje declarativo: especificación de los distintos elementos de configuración (recursos)
- Programado en Ruby
- Permite ser extendida
- Configuración mediante manifiestos:

```
file { ' testfile ':  
  path      => '/tmp/testfile',  
  ensure    => present,  
  mode      => 0640,  
  content   => "I'm a test file .",  
}
```

Infraestructuras de ejecución de trabajos distribuidos

- AppScale
 - Implementación de código abierto del App Engine de Google
 - Alojamiento de aplicaciones web y ejecución de trabajos
- TORQUE
 - Infraestructura clásica de ejecución de trabajos
 - Alojamiento de aplicaciones web y ejecución de trabajos

Infraestructura de servicios web en tres niveles

- Balanceador de carga: distribuir las peticiones web a los servidores web
- Servidores web: procesar peticiones web; lectura o escritura en una base de datos
- Base de datos: almacenamiento de la información

Modelado de recursos distribuidos

- Recursos locales
 - Clásicos en Puppet
 - Locales a un nodo
- Recursos distribuidos
 - No existen en Puppet
 - Características propias:
 - Dependencia
 - Disponibilidad

Modelado de recursos distribuidos

- Tipo
 - Descripción declarativa en la que se definen los atributos del recurso
- Proveedor
 - Implementación necesaria para llevar dicho recurso al estado deseado

Modelado de recursos distribuidos

- Tipo **Recurso distribuido**
 - Nombre
 - Fichero de dominio
 - Conjunto de máquinas físicas
- Proveedor **Recurso distribuido**
 - Puesta en marcha
 - Monitorización
 - Parada

Modelado de recursos distribuidos

- Más detalles del proveedor

