XtreemOS

Enabling Linux
for the Grid

Security and Virtual Organisation Management in XtreemOS

Alvaro Arenas
STFC Rutherford Appleton Laboratory, UK

XtreemOS Summer School, Oxford, September 2010

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Information Society
Technologies

Outline

- **Security Concepts**
- **Grid Security**
  - OGSA Security
  - Grid Security Infrastructure
- **Security and VO Management in XtreemOS**
  - XtreemOS Security Services
  - XtreemOS Trust Model
  - XtreemOS Single Sing-On and Delegation
  - Isolation
- **Concluding Remarks**

Security

- **Computer security deals with the prevention and detection of unauthorised actions by user of a computer system**

- Keep the bad guy out
  - Authentication; firewalls, ...
- Let him in, but keep him from doing damage
  - Access control; sandboxing; ...
- Keep everybody out
  - Isolation; ...
- Catch him and prosecute him
  - Monitoring; auditing; ...

Security in Operating Systems

- **Identification and authentication**
  - Be sure about the identity of the user
- **Process management**
  - Protect one process from another
- **Memory management**
  - Protect the memory of one process from another
- **File management**
  - Protect the files owned by each user
- **Audit controls**
  - Log security-sensitive operations and report them to administrators
- **Recovery**
  - Allow system recovery if security breach occurs

Security in Grids

- **Grids concern with …**
  - "Coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organisations."
    - From the "*Anatomy of the Grid*"

- **So Grid Security is security to enable Virtual Organisationss**
  - Access to shared services/resources
  - Cross-domain authentication, authorisation, accounting, billing
  - May contain individuals acting alone – their home organisation administration need not necessarily know about all activities
  - Leave resource owner always in control

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Security in a Grid OS

- **Native support for VO management**
  - XtreemOS embeds VO-management functionalities into the Linux kernel
- **Leverage OS security support to protect resources**
  - XtreemOS exploits OS isolation capabilities (Linux containers) to provide strong isolation and fine-grained control of resource usage
  - Map VOs policies into access control policies
- **Transparent security management**
  - Flexible management of certificates, making its operation as transparent as possible for end users
- **Scalability in security**
  - Separate resource management from VO and user management

Outline

- Security Concepts

- Grid Security
  - OGSA Security
  - Grid Security Infrastructure

- Security and VO Management in XtreemOS
  - XtreemOS Security Services
  - XtreemOS Trust Model
  - XtreemOS Single Sing-On and Delegation
  - Isolation

- Concluding Remarks

Virtual Organisations

- **A VO is**
  - a temporary or permanent coalition of geographically dispersed and autonomous participants
    - including individual and/or organisations,
  - who agree to share resources in the system in order to fulfill their tasks
    - e.g. running jobs, sharing applications, accessing data

- **Properties**
  - Geographically distributed
  - Autonomously governed
  - Short-termed or long-term
  - Static or dynamics

Information Society
Technologies

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Virtual Organisations and Trust

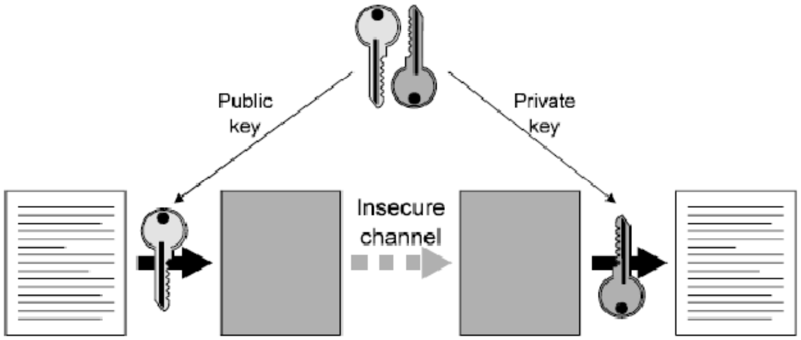
- **VOs are used as a bridge to provide a Grid security solution based on trust**
  - The extent to which a participant can rely on others to behave

- **Establishing trust**
  - Personal recommendations
  - Reputation from trusted sources
  - Cryptographic verification of the information given

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Trust through Cryptography

- **An entity uses computer programs to cryptographically verify the information given**
  - If everything is ok, then trust of the information is established
  - Otherwise, there is not trust



- **Public key encryption**
  - Users possess public/private key pairs
  - Anyone can encrypt with the public key, only one person can decrypt with the private key

Information Society Technologies
XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Certification Authorities (CAs)

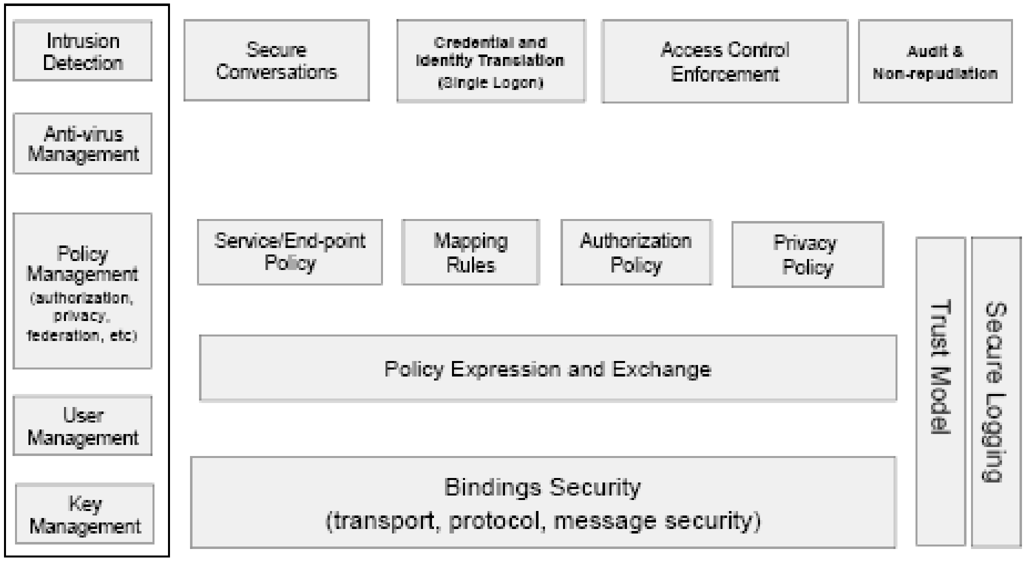
- The CAs are responsible for certifying the public keys of different users who subscribe to the CA
  - Guarantee the connection between a key and an end entity
- CAs are entities that are trusted by different systems
- An end entity is
  - Person, role (“Director of marketing”), organisation, pseudonym, a piece of hardware or software, an account (bank or credit card)
- CA manages key lifecycle: creation, store, delete, renew

OGSA Security

- **Secure functionality should be cast as services**
  - allowing applications to locate and use the particular functionality they need

- **Leverage on existing and emerging WS security standards**
  - Authentication service;
  - Identity mapping service
  - Authorisation service;
  - VO Policy service;
  - Credential conversion service;
  - Audit service; etc

Components of the Grid Security Model



Intrusion Detection

Anti-virus Management

Policy Management (authorization, privacy, federation, etc)

User Management

Key Management

Secure Conversations

Credential and Identity Translation (Single Logon)

Access Control Enforcement

Audit & Non-repudiation

Service/End-point Policy

Mapping Rules

Authorization Policy

Privacy Policy

Policy Expression and Exchange

Bindings Security
(transport, protocol, message security)

Trust Model

Secure Logging

GSI – Grid Security Infrastructure

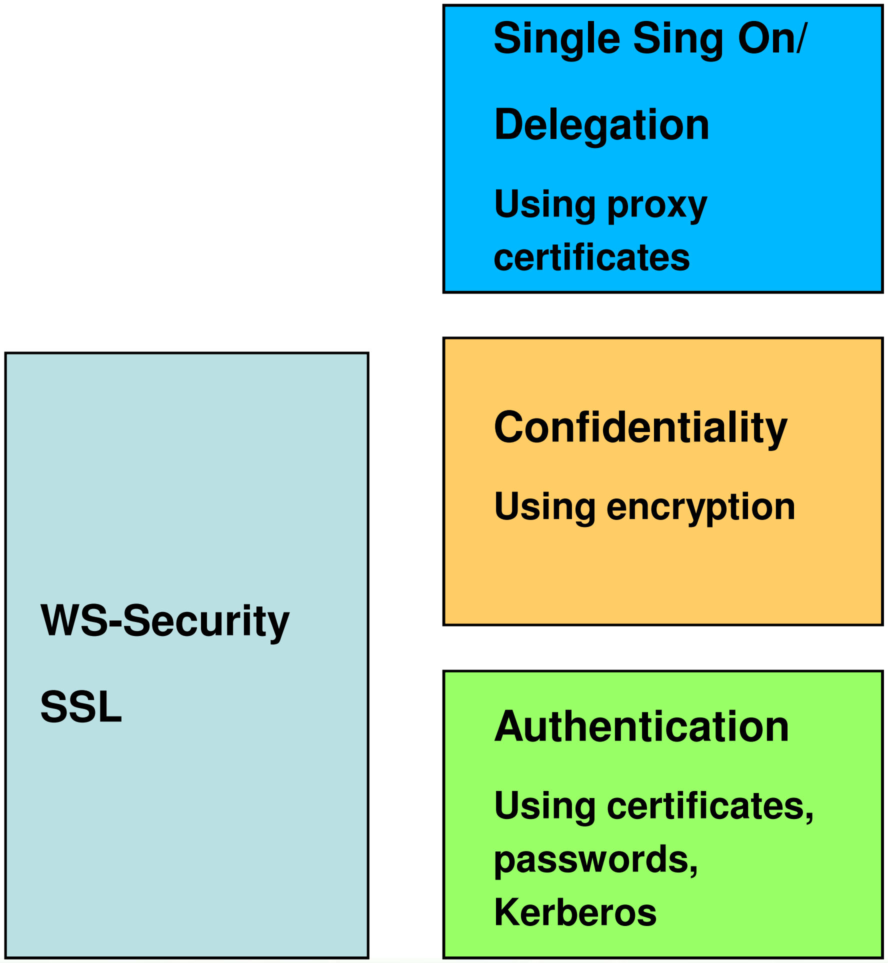
- A reference specification for Grid security architectures
- Protocols and APIs to address Grid security needs
- Based on public-key encryption technology
    - SSL protocol for authentication, message protection
    - X.509 certificates
- Each user as a Grid id, a private key, and a certificate signed by a CA

- First implementation – in the Globus Toolkit

Information Society Technologies

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

High-Level View of GSI



Single Sing On/

Delegation

Using proxy certificates

Confidentiality

Using encryption

Authentication

Using certificates, passwords, Kerberos

WS-Security

SSL

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Authentication in GSI

- Certificate-based authentication (PKI)
- GSI certificate includes information such as
  - Subject name;
  - public key belonging to the subject;
  - Identity of the CA; and
  - Digital signature of the named CA
- Certificates are obtained via established protocols

Single Sign On and Delegation

- **Jobs require access to multiple resources**
  - To authenticate with your certificate directly you would have to type a passphrase every time
- **Need to automate access to other resources: Authenticate Once**
  - Important for complex applications that need to use Grid resources
  - Allows remote processes and resources to act on user's behalf - also known as **delegation**
  - Also you need a way to send you VO details (Groups membership, roles and capabilities) across
- **Solution adopted in the GSI: proxy certificates**
  - A temporary key pair
  - in a temporary certificate signed by your 'long term' private key
  - valid for a limited time (default: 12 hours), but can be renewed

Outline

- Security Concepts

- Grid Security
  - OGSA Security
  - Grid Security Infrastructure

- Security and VO Management in XtreemOS
  - XtreemOS Security Services
  - XtreemOS Trust Model
  - XtreemOS Single Sing-On and Delegation
  - Isolation

- Concluding Remarks

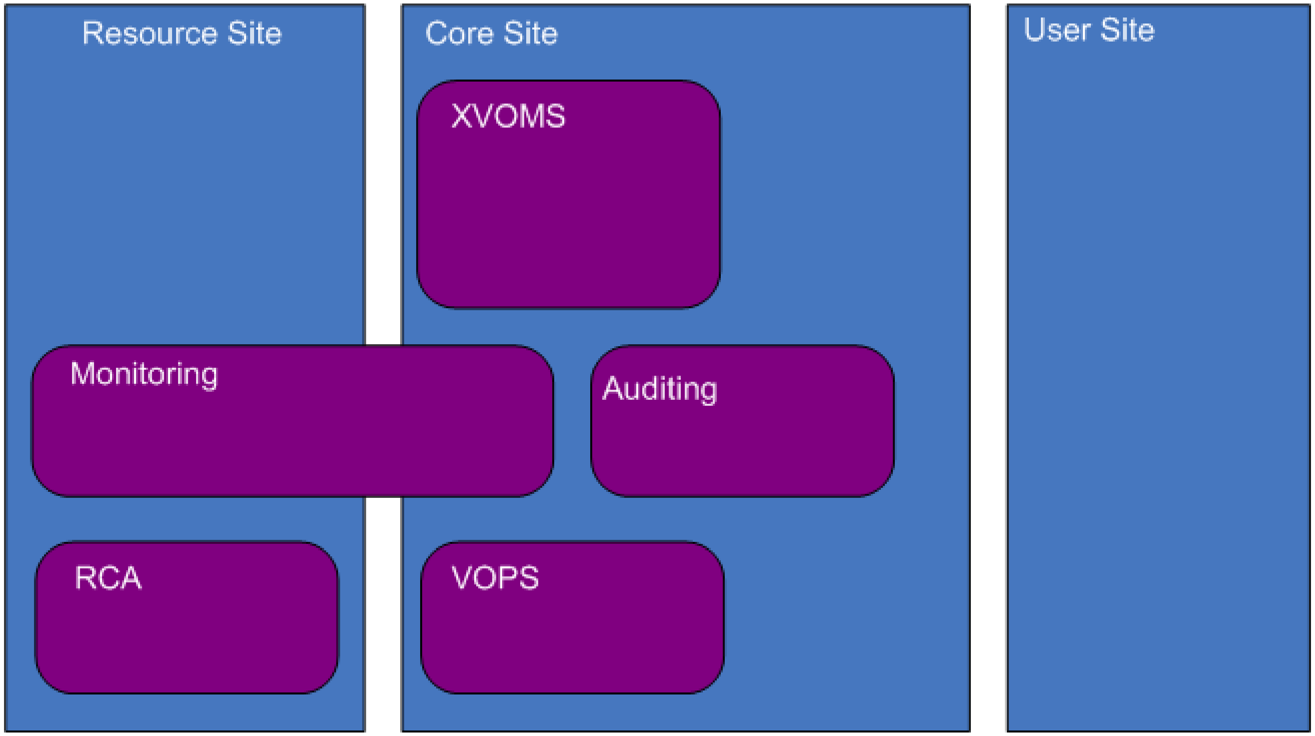
XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

XtreemOS System

- **A XtreemOS system consists of**
  - A set of **resource machines** from one or more participants
    - Offering resources through a set of foundation-level node services
  - A set of **Grid-wide system services**
  - A set of **VOs** to support cross-machine resource sharing and logical isolation of resource usage within the system

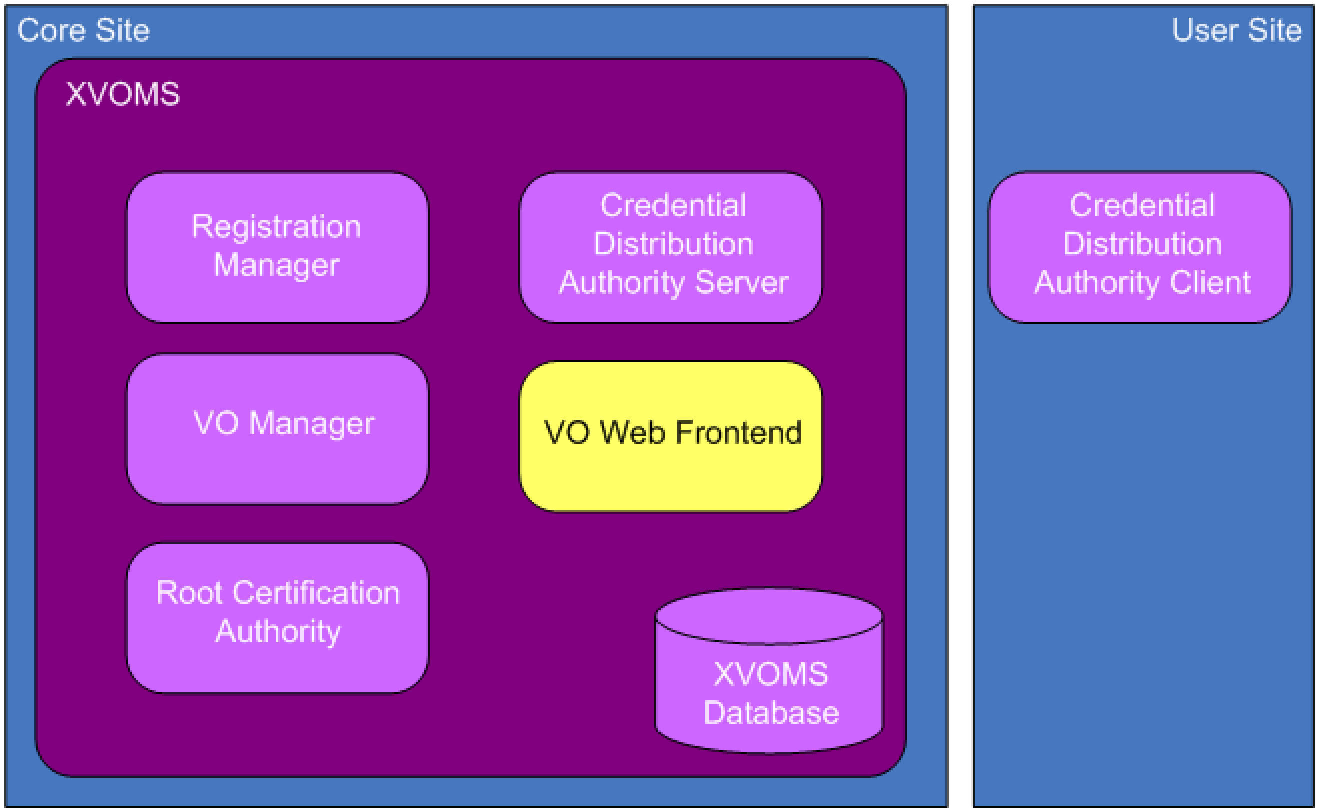
- **A user of a XtreemOS system is defined as another system**
  - Including humans or separated autonomous software systems
  - Interacts with the current system through a set of well-defined interfaces.

VO Management and Security Services

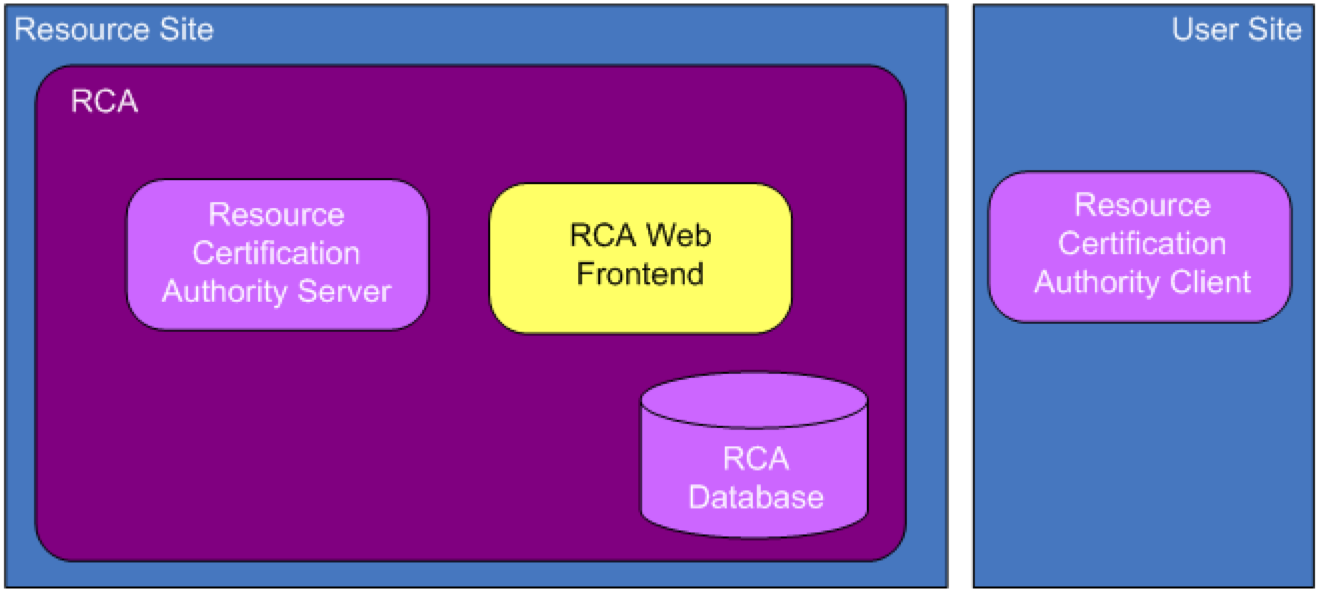


XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

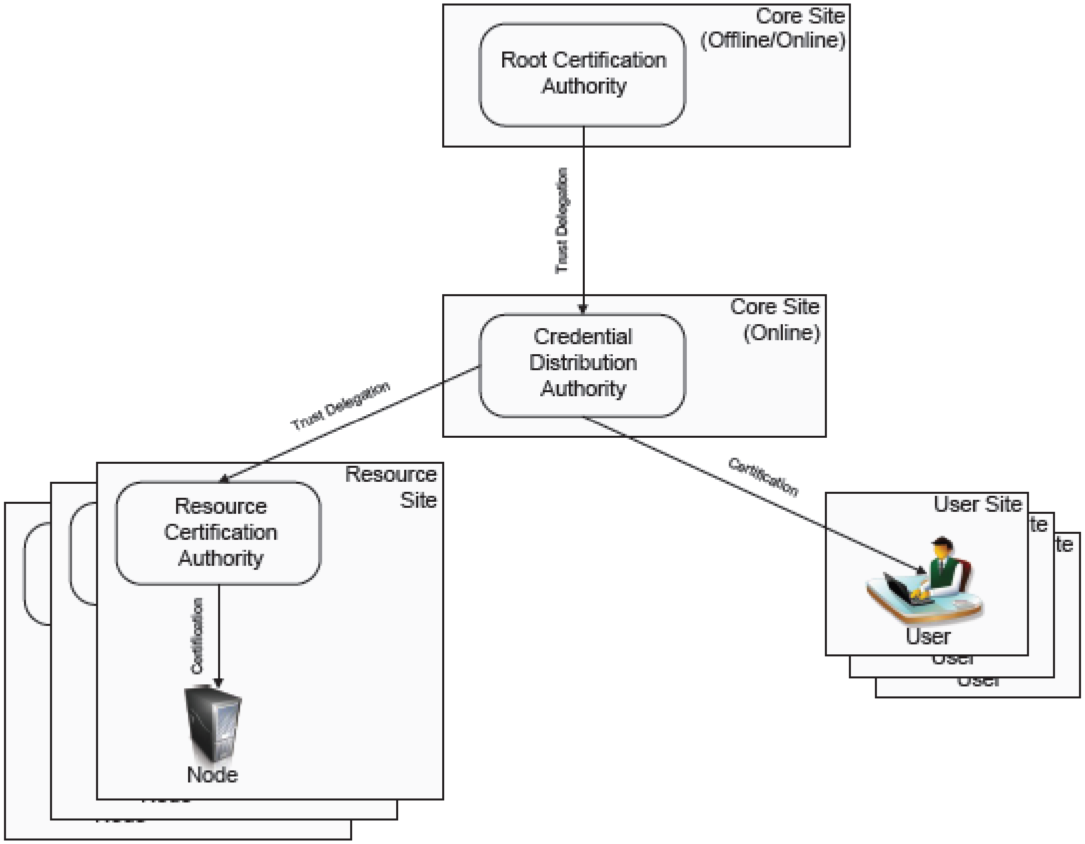
Virtual Organisation Manager Service (X-VOMS)


Core Site
XVOMS
Registration Manager
Credential Distribution Authority Server
VO Manager
VO Web Frontend
Root Certification Authority
XVOMS Database
User Site
Credential Distribution Authority Client

Resource Certification Authority (RCA)


Resource Site
RCA
Resource Certification Authority Server
RCA Web Frontend
RCA Database
User Site
Resource Certification Authority Client

XtreemOS Trust Model



Building Up Trust in XtreemOS

- **XVOMS Certificate**
  - XVOMS has a self-signed certificate representing the root certificate of the system
  - The private counterpart is used by CDA to sign end-entity certificates for users and subordinated RCAs
- **User registration with XVOMS**
  - Each user shares a secret (i.e. password) with XVOMS
  - User obtain XVOMS public key certificate through established password-based mutual authentication protocols
  - There is not need of pre-installed certificate
- **RCA registration with XVOMS**
  - Each RCA is registered with a XVOMS and is given a shared secret with XVOMS
  - Mutually authenticate with XVOMS with any pre-installed certificate
- **Machine registration with RCA**

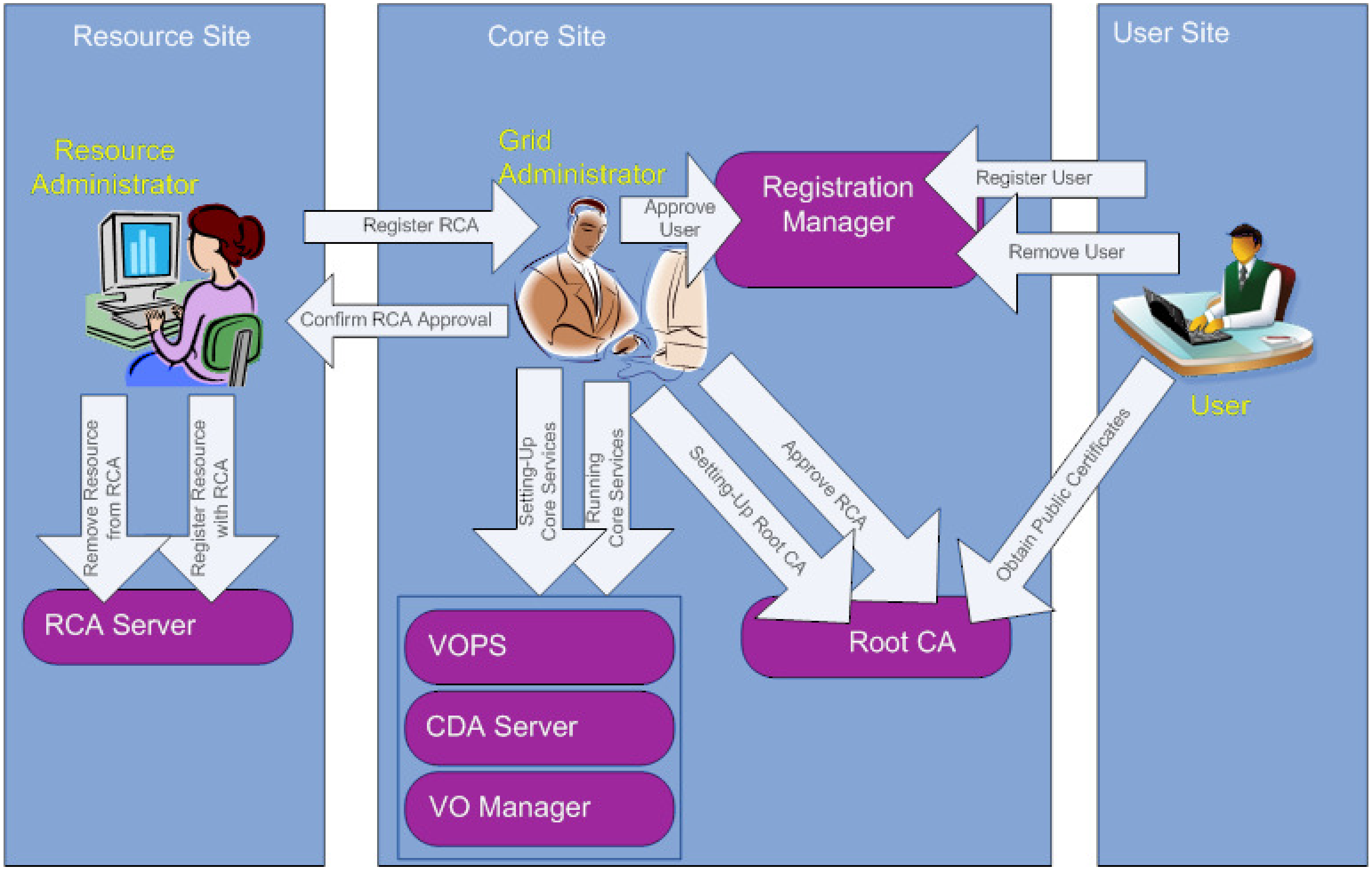
Advantages of XtreemOS Trust Model

- **User management is separated from resources management**
- **Scalability in resource management**

- **Main difference with classical PKI trust models resides in the set up of trust**
  - In classical PKI models, trusted root CA certificates are distributed through offline means
  - In XtreemOS, certificates could be created on-the-fly and disseminated through online protocols

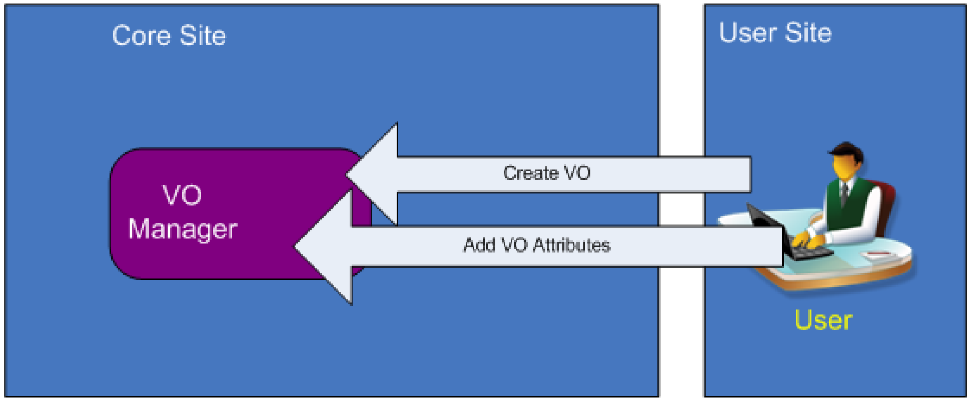
- **SSO and Delegation**
  - Not depending on proxy certificates

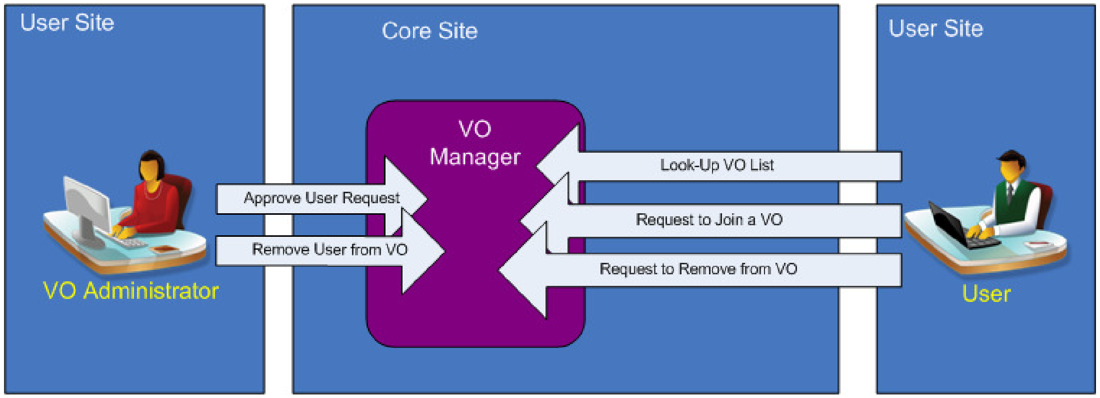
Grid Management Capabilities



XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

VO Management Capabilities (User Site)

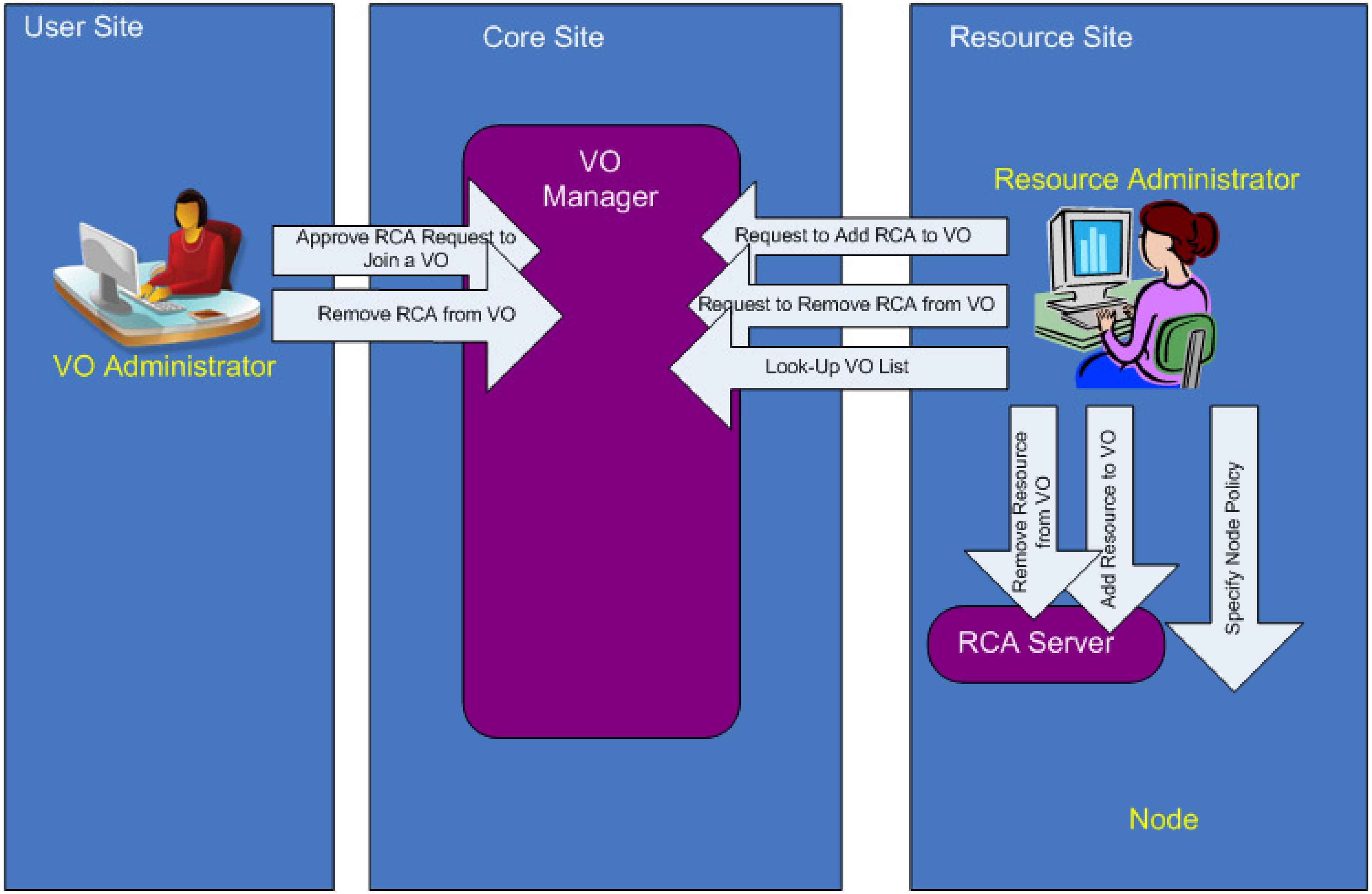




Information Society Technologies

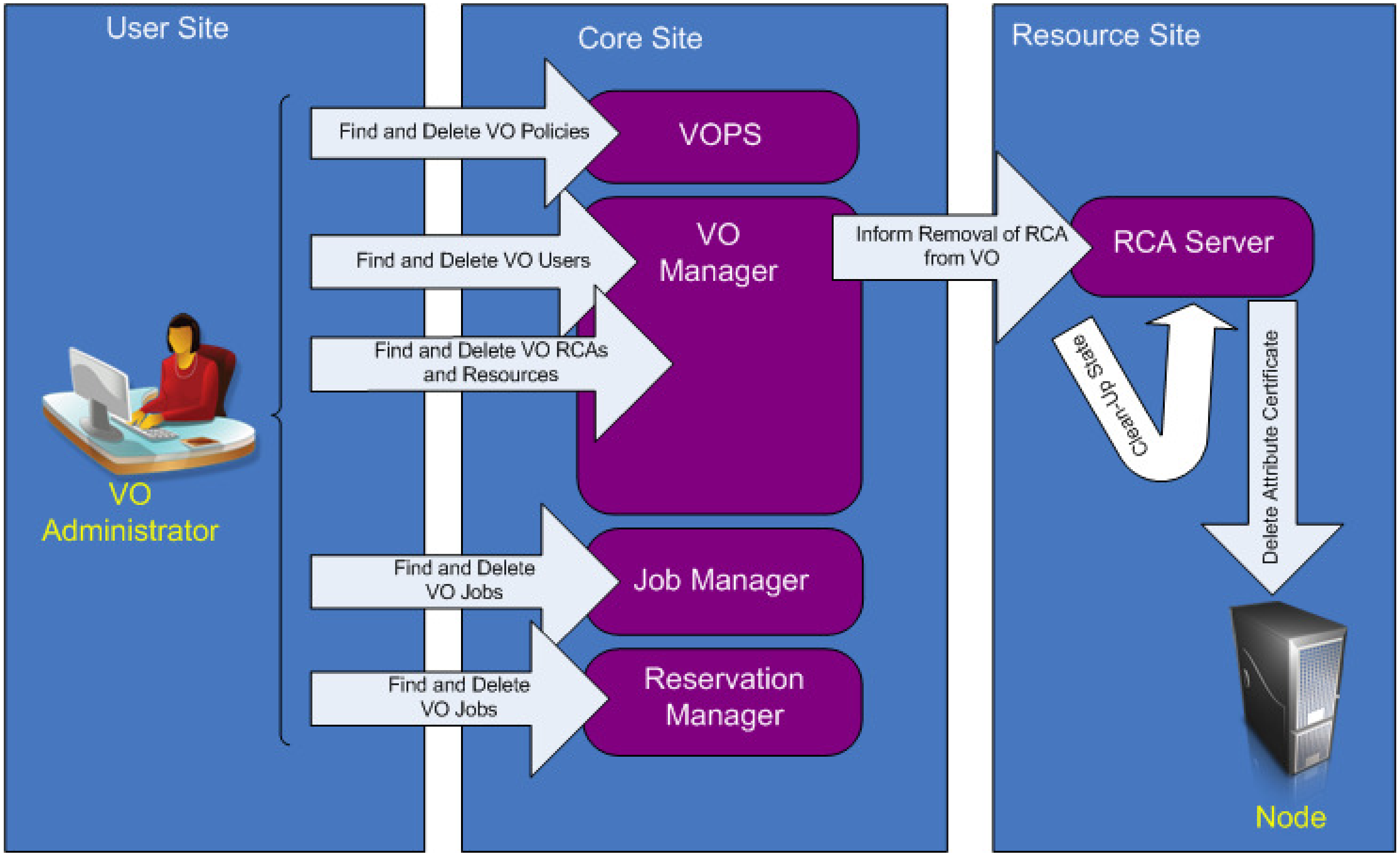
XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

VO Management Capabilities (Resource Site)



XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

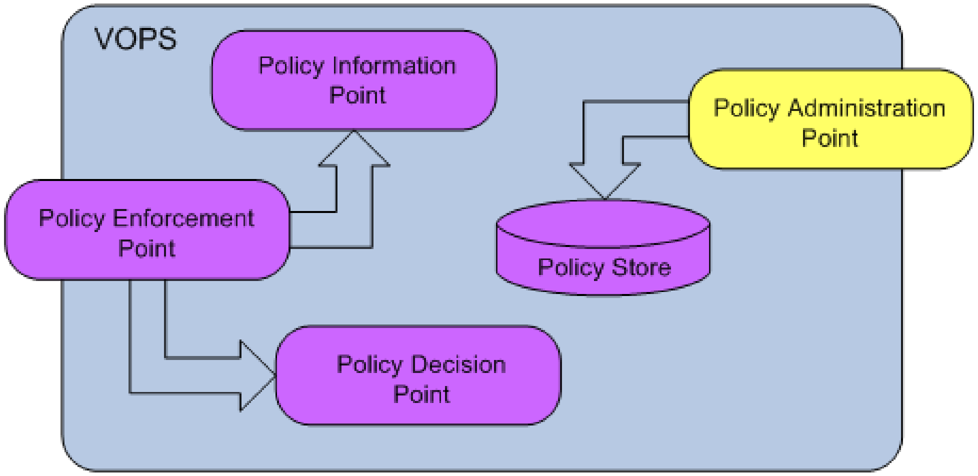
VO Management Capabilities (VO Termination Phase)



XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

Policy Management in XtreemOS

- **Dealing with four type of policies**
  - User; resource;
  - VO; and filtering policies
- **XACML as policy language**
- **Policies are evaluated at**
  - **Selection time**: to ensure that resources selected are suitable
  - Access time: to control access to resources



XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576

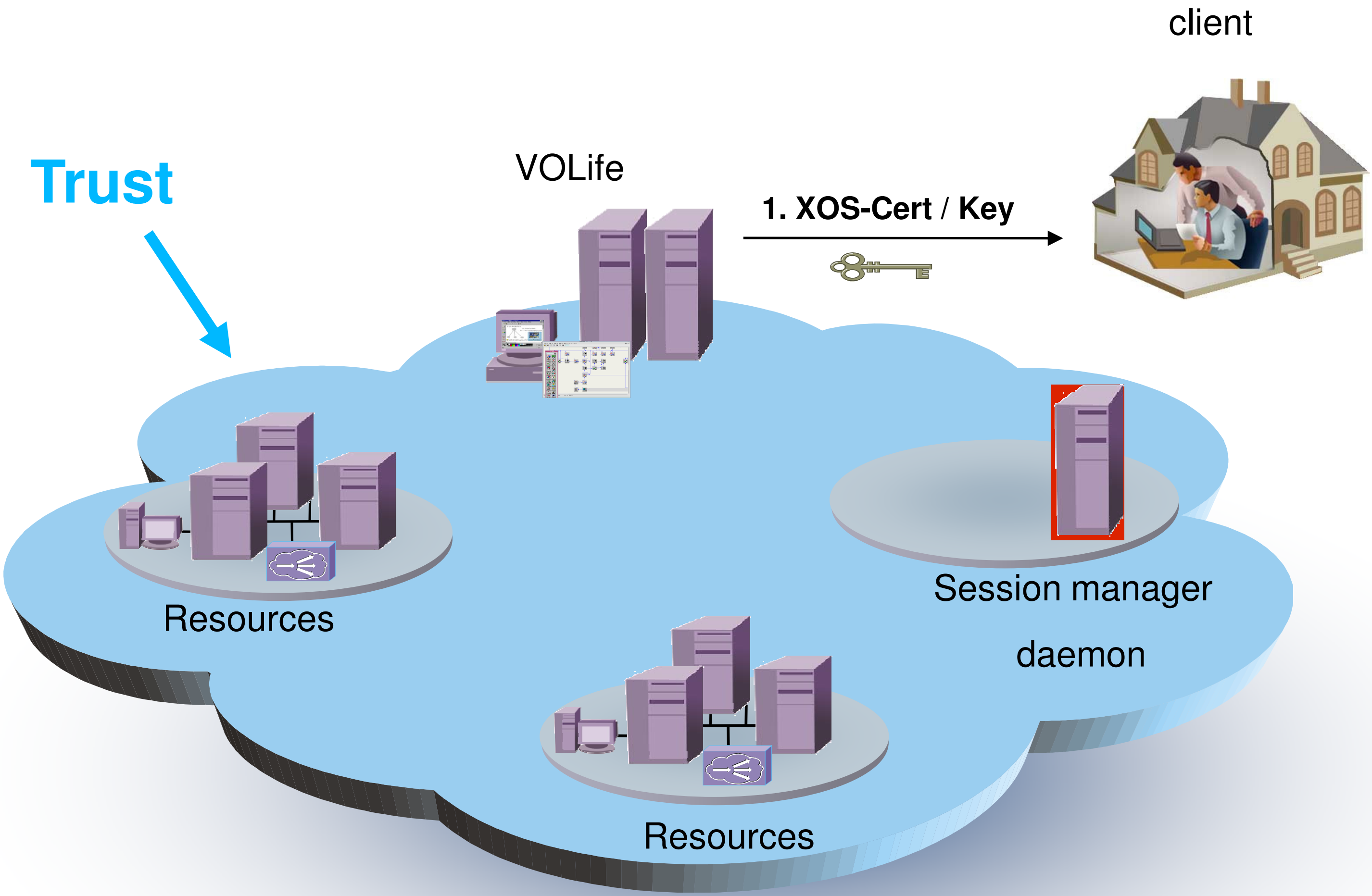
Single Sign-On and Delegation

- **Single Sing-On**
  - As a distributed OS, XtreemOS services trust each other
  - Once user credentials are validated by a XtreemOS service, they can be used by other XtreemOS services without additional validation

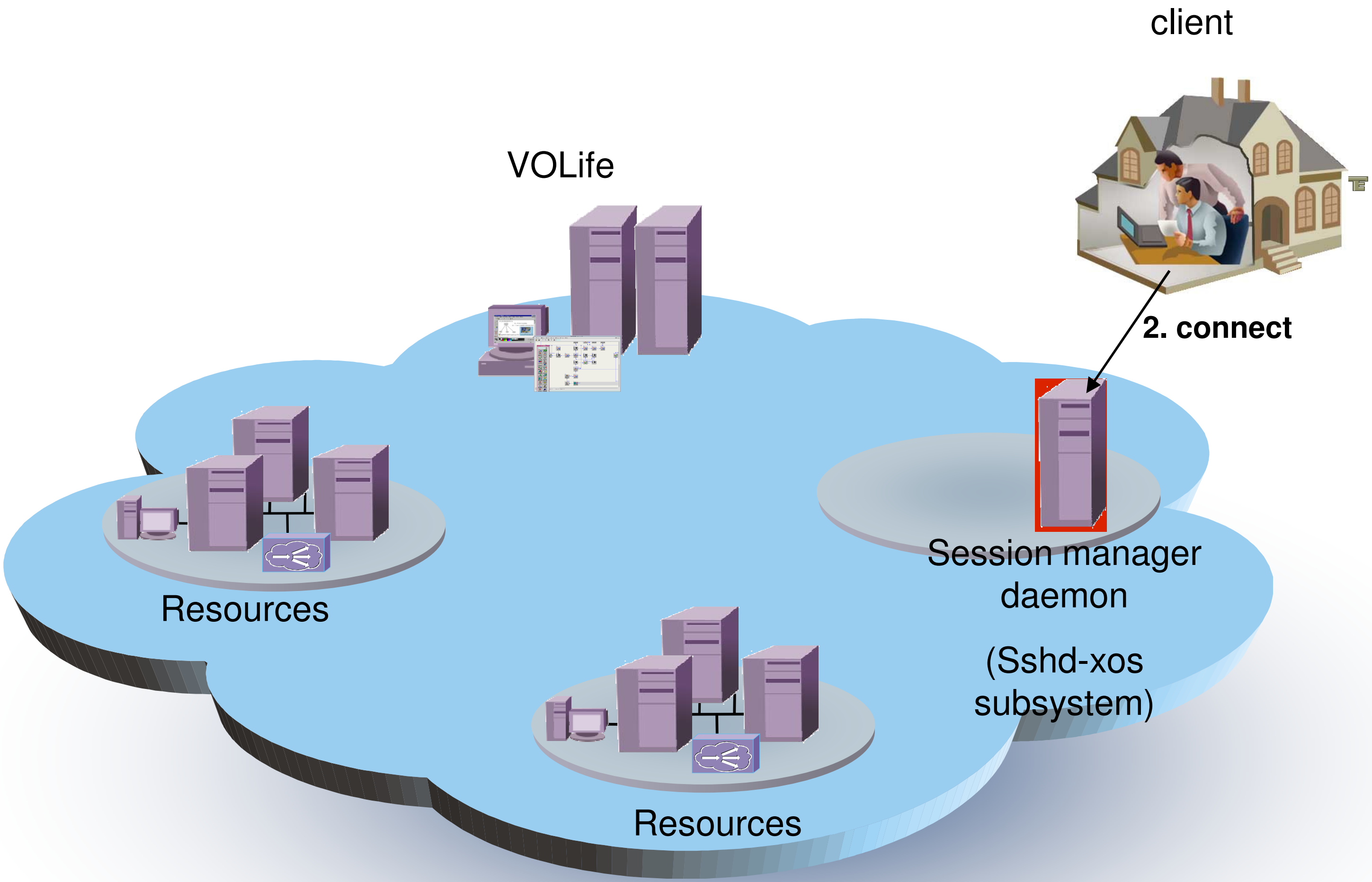
- **Underpinning technology**
  - A trusted credential store service is associate to each user session.
    - Authenticate the user when he opens a session,
    - Collect and validate all user credentials,
    - Forward all grid requests (xsub, xps, etc.) from the user to XtreemOS services
  - **There is not need of proxy certificates**

- **Delegation, exploiting similar technology**
  - A credential store services is associated to jobs on the same resource node
  - Once job credentials are validated, they can be used in other XtreemOS services
  - Key technology for interactive jobs

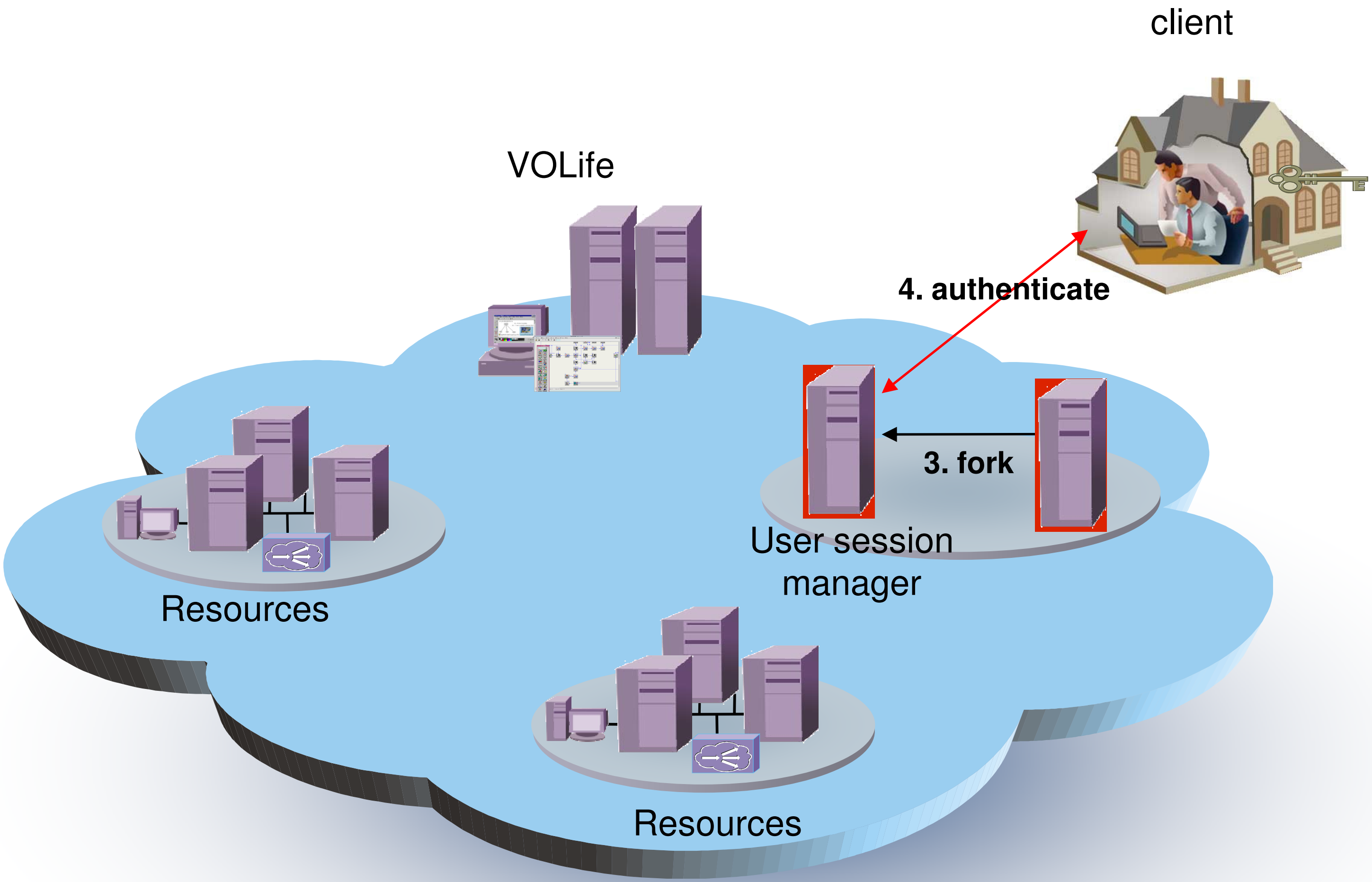
Single-sign-on for grid requests



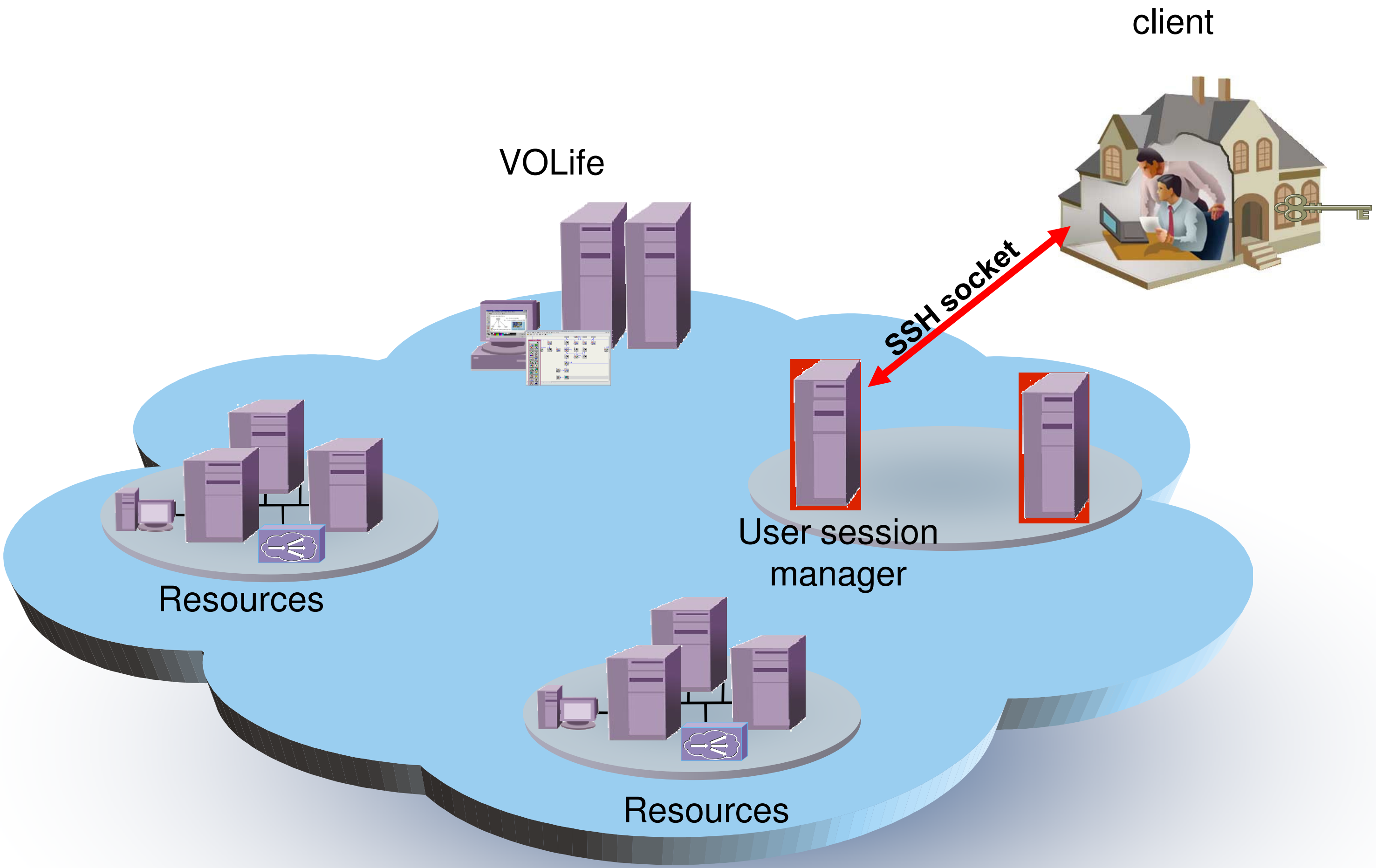
Session managers



Login

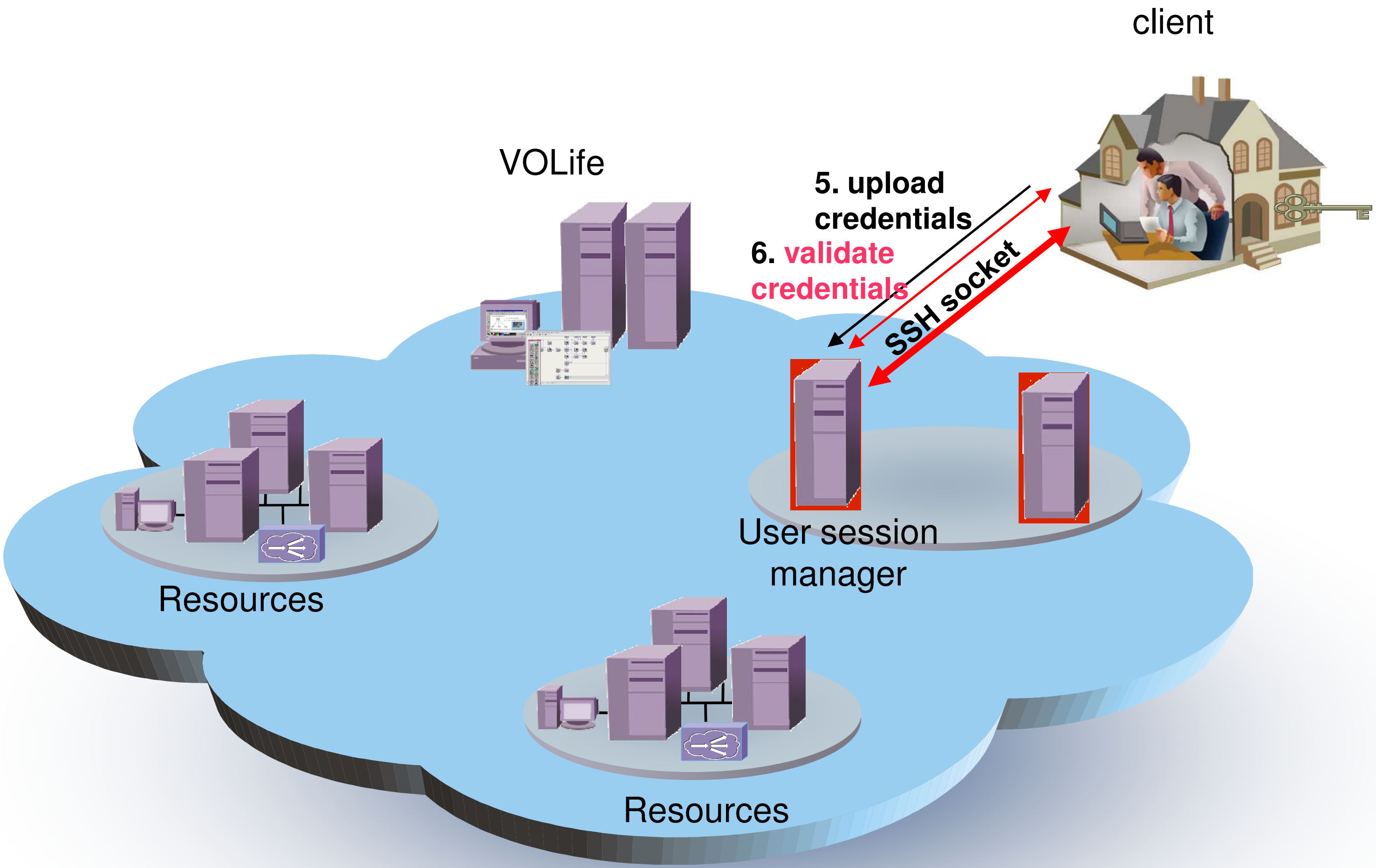

client
VOLife
4. authenticate
3. fork
User session manager
Resources
Resources

SSH-XOS control master



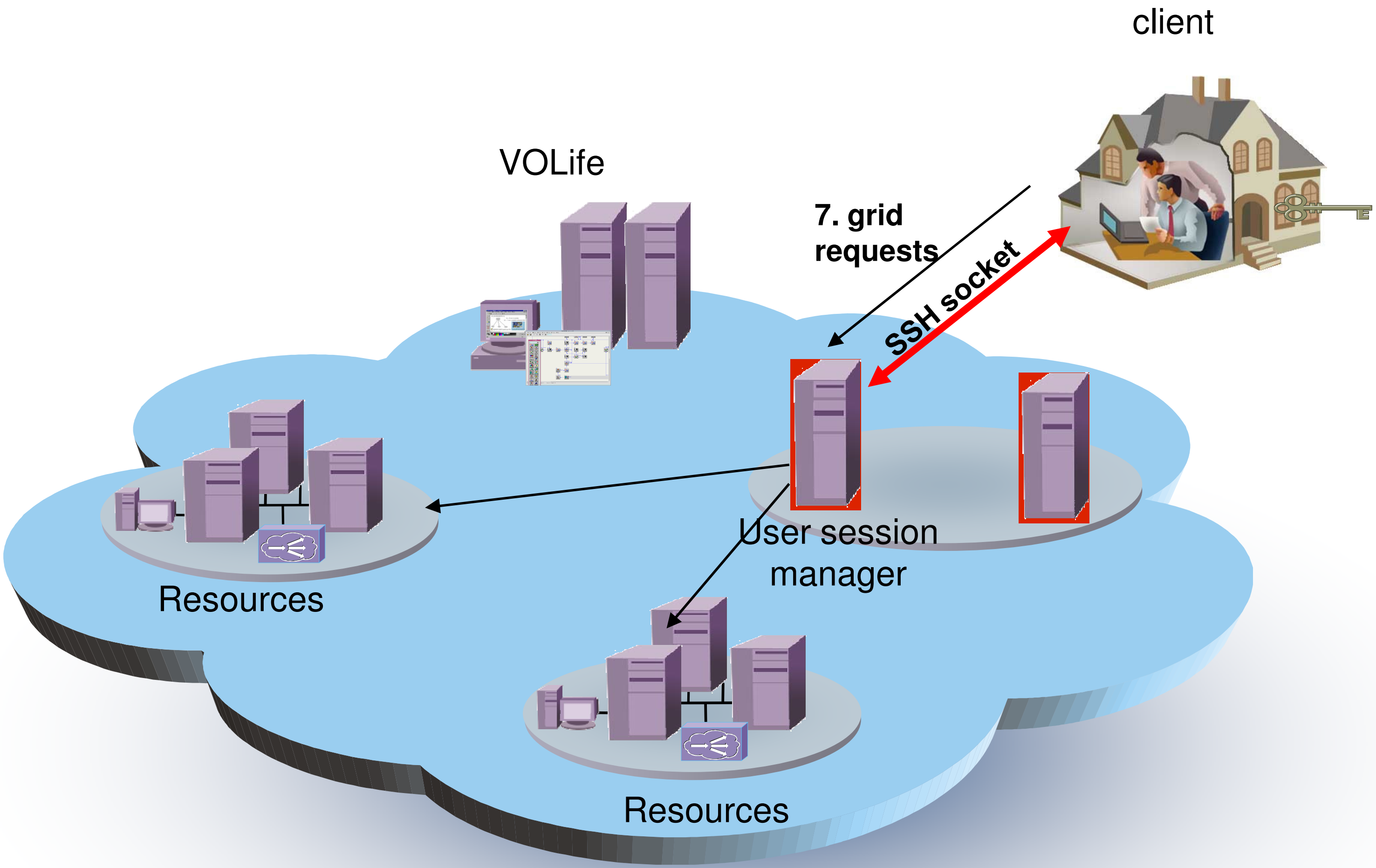
XtreemOS
Enabling Linux for the Grid

Credential management



Information Society Technologies

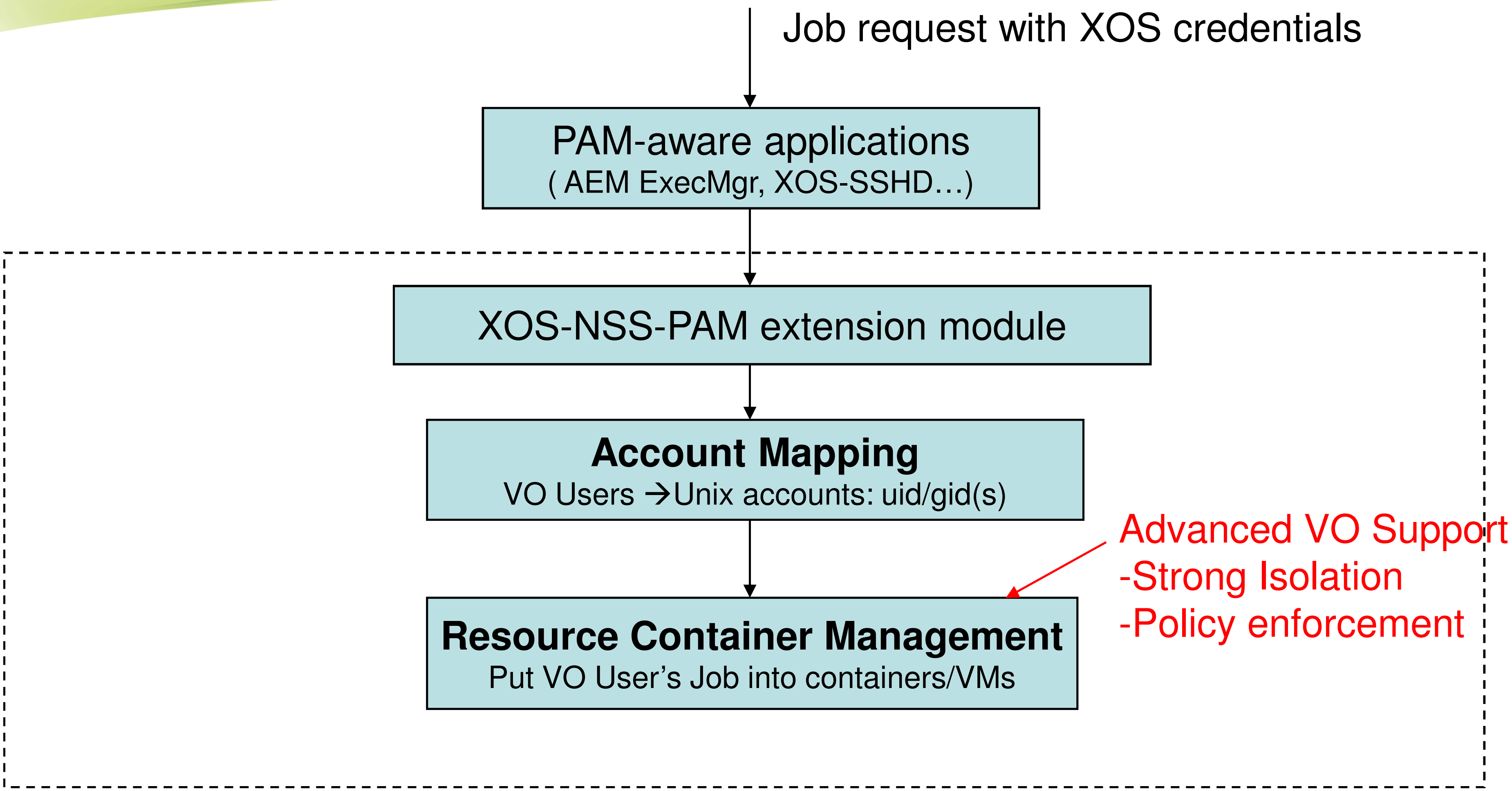
Grid requests submission



Isolation in XtreemOS

- **Basic idea: Put each job (PAM session) into a resource container**
  - A resource container could be seen as a virtual machines in a local OS instance
  - A resource container allows **fine-grained**, **isolated** and **strong** control of resource usage of a job (could be a hierarchy of processes)
- **Features: Full-fledged control of resource usage by VOs**
  - CPU: Assignment of cores, bandwidth/percentage/ priority/walltime allocation
  - Memory: virtual/physical/swap memory limitation
  - Disk I/O: disk i/o bandwidth limitation
  - Network: network bandwidth/traffic limitation

Isolation



Job request with XOS credentials

PAM-aware applications
(AEM ExecMgr, XOS-SSHD…)

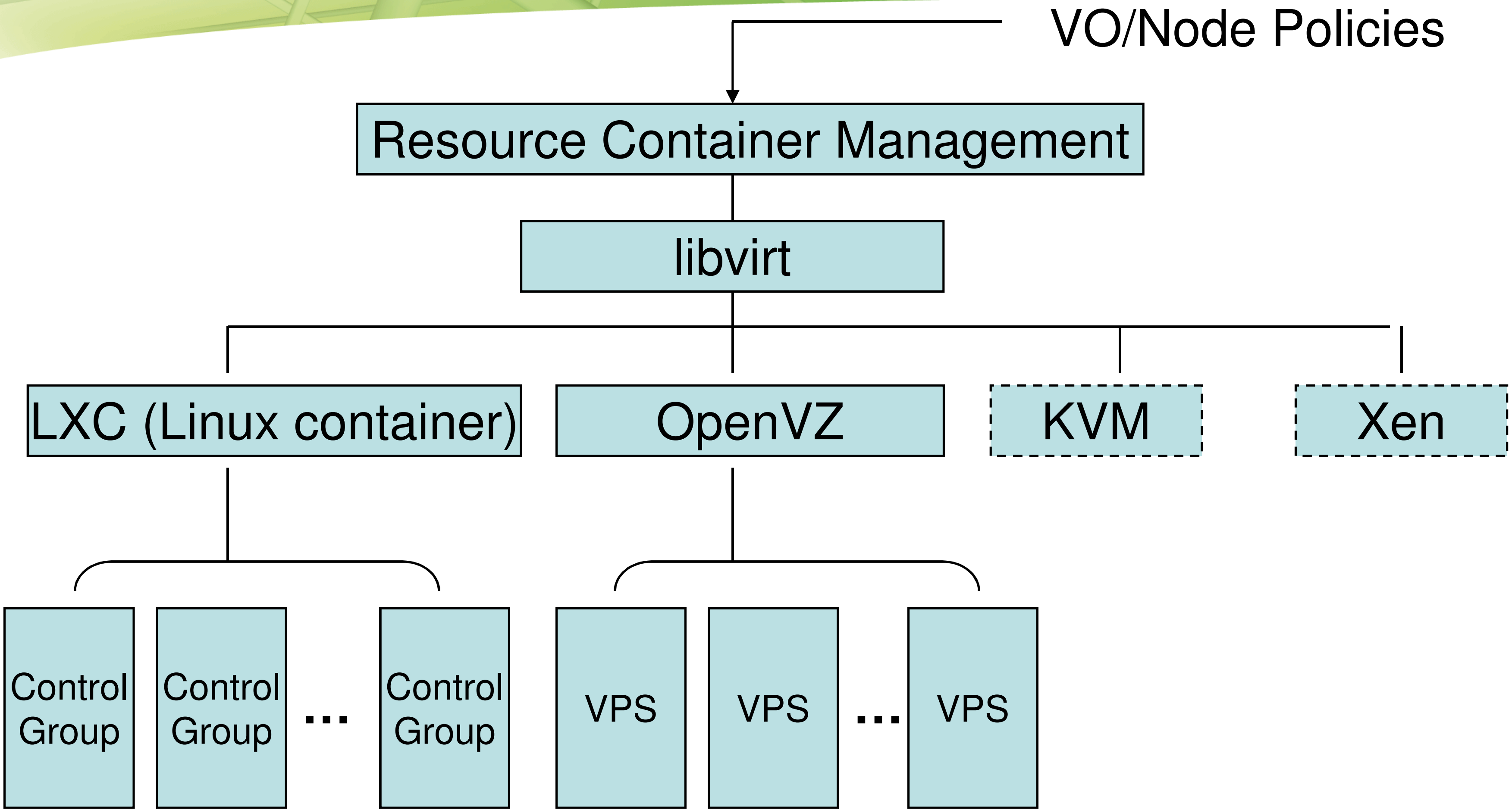
XOS-NSS-PAM extension module

Account Mapping
VO Users →Unix accounts: uid/gid(s)

Resource Container Management
Put VO User's Job into containers/VMs

Advanced VO Support
-Strong Isolation
-Policy enforcement

How to do it ?



VO/Node Policies

Resource Container Management

libvirt

LXC (Linux container)

OpenVZ

KVM

Xen

Control Group

Control Group

…

Control Group

VPS

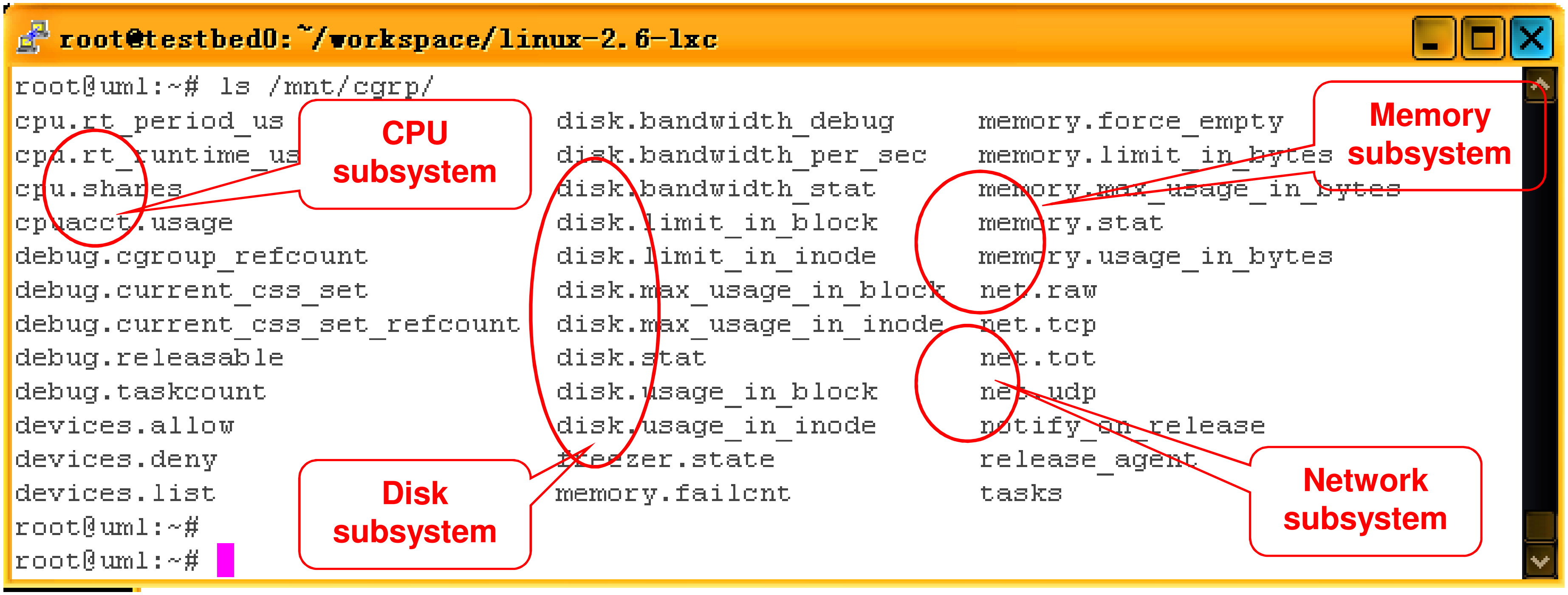
VPS

…

VPS

What we achieve now?

- In advanced version of VO-support, what new features have been embedded in based on cgroup mechanism?



```
root@uml:~# ls /mnt/cgrp/
cpu.rt_period_us                disk.bandwidth_debug     memory.force_empty
cpu.rt_runtime_us               disk.bandwidth_per_sec   memory.limit_in_bytes
cpu.shares                      disk.bandwidth_stat      memory.max_usage_in_bytes
cpuacct.usage                   disk.limit_in_block      memory.stat
debug.cgroup_refcount           disk.limit_in_inode      memory.usage_in_bytes
debug.current_css_set           disk.max_usage_in_block  net.raw
debug.current_css_set_refcount  disk.max_usage_in_inode  net.tcp
debug.releasable                disk.stat                net.tot
debug.taskcount                 disk.usage_in_block      net.udp
devices.allow                   disk.usage_in_inode      notify_on_release
devices.deny                    freezer.state            release_agent
devices.list                    memory.failcnt           tasks
root@uml:~#
root@uml:~#
```

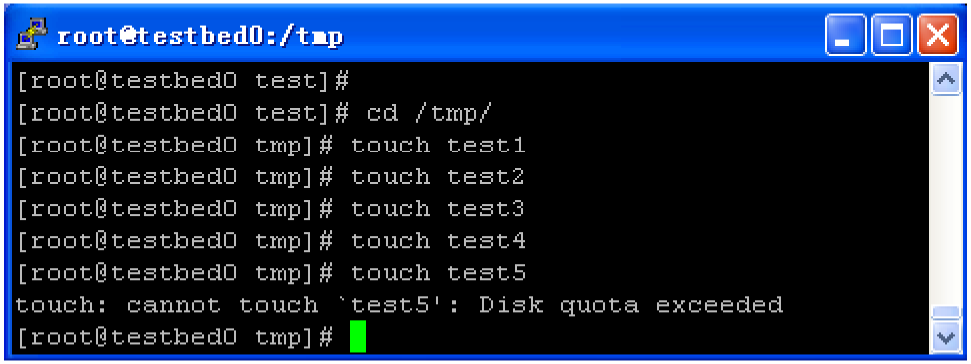
Snapshot of subsystem functionalities

Disk quota limitation

- Record the usage of allocated file inode
- Record the usage of allocated disk block

- Limit created file number

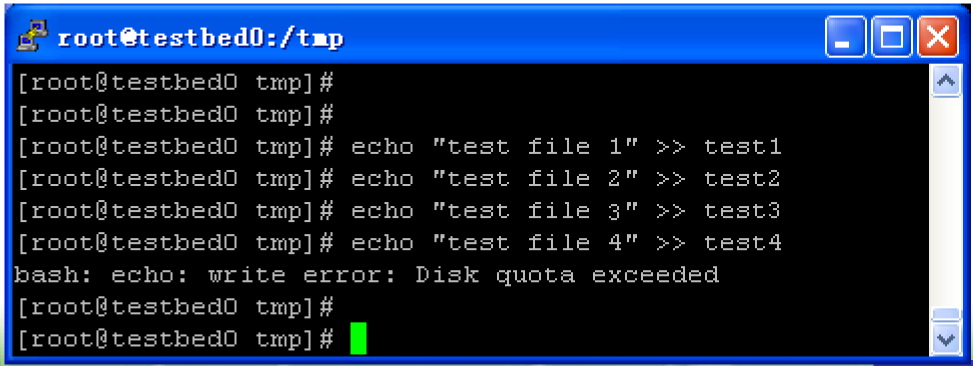
# echo 4 > disk.max_usage_in_inode



```
[root@testbed0 test]#
[root@testbed0 test]# cd /tmp/
[root@testbed0 tmp]# touch test1
[root@testbed0 tmp]# touch test2
[root@testbed0 tmp]# touch test3
[root@testbed0 tmp]# touch test4
[root@testbed0 tmp]# touch test5
touch: cannot touch `test5': Disk quota exceeded
[root@testbed0 tmp]#
```

- Limit allocated file block (3*4096)

# echo 1288 > disk.max_usage_in_block



```
[root@testbed0 tmp]#
[root@testbed0 tmp]#
[root@testbed0 tmp]# echo "test file 1" >> test1
[root@testbed0 tmp]# echo "test file 2" >> test2
[root@testbed0 tmp]# echo "test file 3" >> test3
[root@testbed0 tmp]# echo "test file 4" >> test4
bash: echo: write error: Disk quota exceeded
[root@testbed0 tmp]#
[root@testbed0 tmp]#
```

Outline

- Security Concepts

- Grid Security
  - OGSA Security
  - Grid Security Infrastructure

- Security and VO Management in XtreemOS
  - XtreemOS Security Services
  - XtreemOS Trust Model
  - XtreemOS Single Sing-On and Delegation
  - Isolation

- **Concluding Remarks**

Security in XtreemOS

- Scalable VO management
  - Independent user and resource management
  - Interoperability with VO management frameworks and security models
  - Customizable isolation, access control and auditing
- Very Dynamic VOs
  - Short-lived VOs created automatically for the duration of an application/workflow
    - Multi-users
  - Lightweight configuration of resources
  - Predefined policies (VO-based)

Security in XtreemOS

- Improving usability
  - Local resource administrator: autonomous management of local resources
  - VO administrator: flexibility management of credential and VO policies
  - End user: login as a Grid user into a VO; the Grid should be as much as possible invisible
- Secure and reliable application execution
  - Fine-grained control of resource usage

On-going and Future Work

- **Traceability**
  - Exploiting tokens for traceability in SSO
- **Security monitoring and auditing**
  - Rule-based monitoring systems; including aggregation of events and logs for auditing purpose
- **Interoperability by using third-party identity providers**
  - Shibboleth; myProxy
- **Evaluating how to adapt some services for the Cloud**
  - Identity as a service

Acknowledgments

- This work is a summary of the work carried out in XtreemOS WP2.1 and WP3.5 work packages

- **INRIA:** Christine Morin, Yvon Jegou
- **ICT:** Haiyan Yu
- **SAP:** Philip Robinson
- **STFC:** Benjamin Aziz, Ian Johnson, Brian Matthews, Erica Yang
- **XLAB:** Matej Artac

XtreemOS

Enabling Linux
for the Grid

Security and Virtual Organisation Management in XtreemOS

Alvaro Arenas

STFC Rutherford Appleton Laboratory, UK

XtreemOS Summer School, Oxford, September 2010

Information Society
Technologies

XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576