



XtreemOS: a Linux-based Operating System for Large Scale Dynamic Grids

Christine Morin

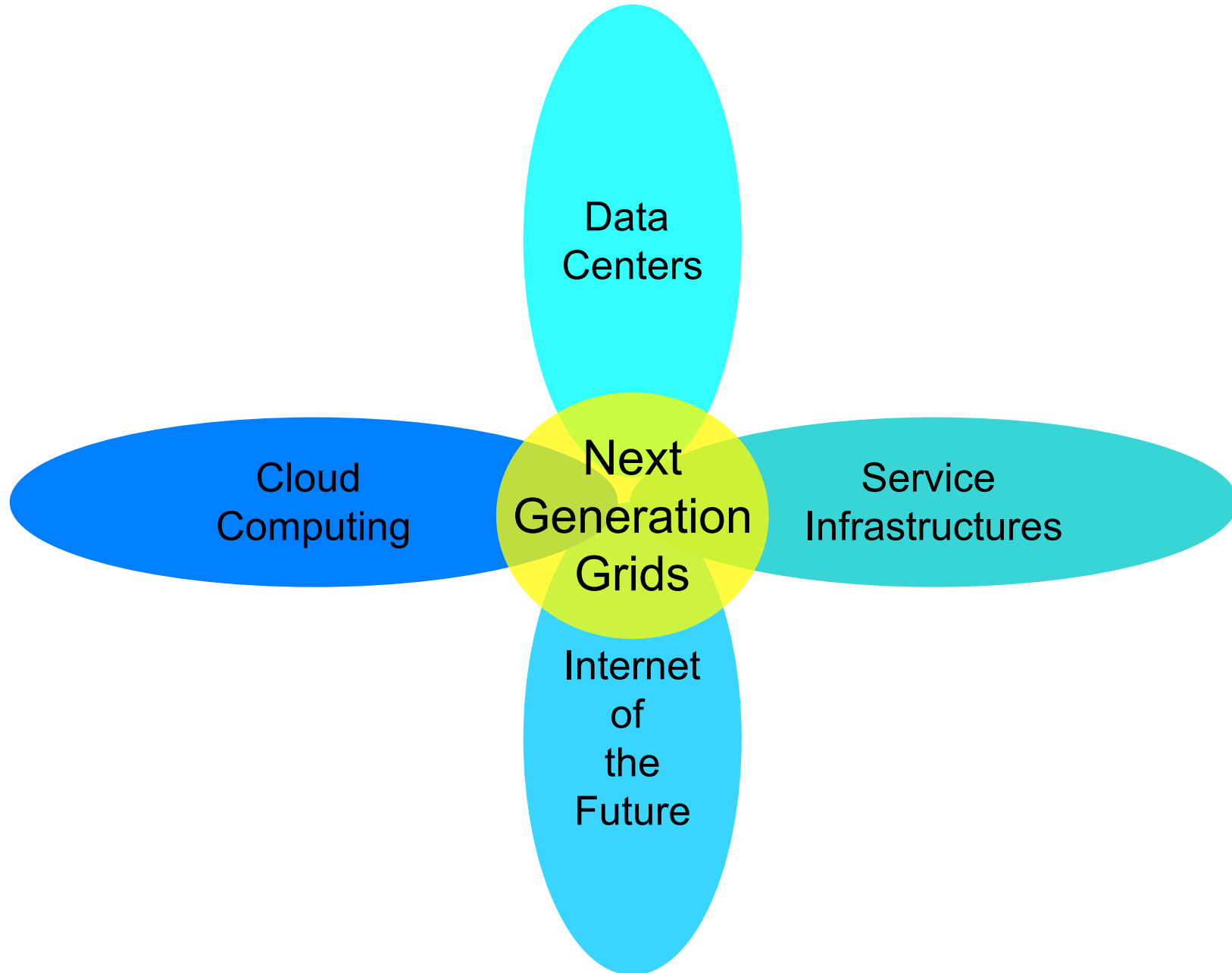
XtreemOS Scientific Coordinator

INRIA Rennes - Bretagne Atlantique

contact@xtreemos.eu

What is XtremOS?

Linux-based Operating System
with native Virtual Organization support
for Next Generation Grids

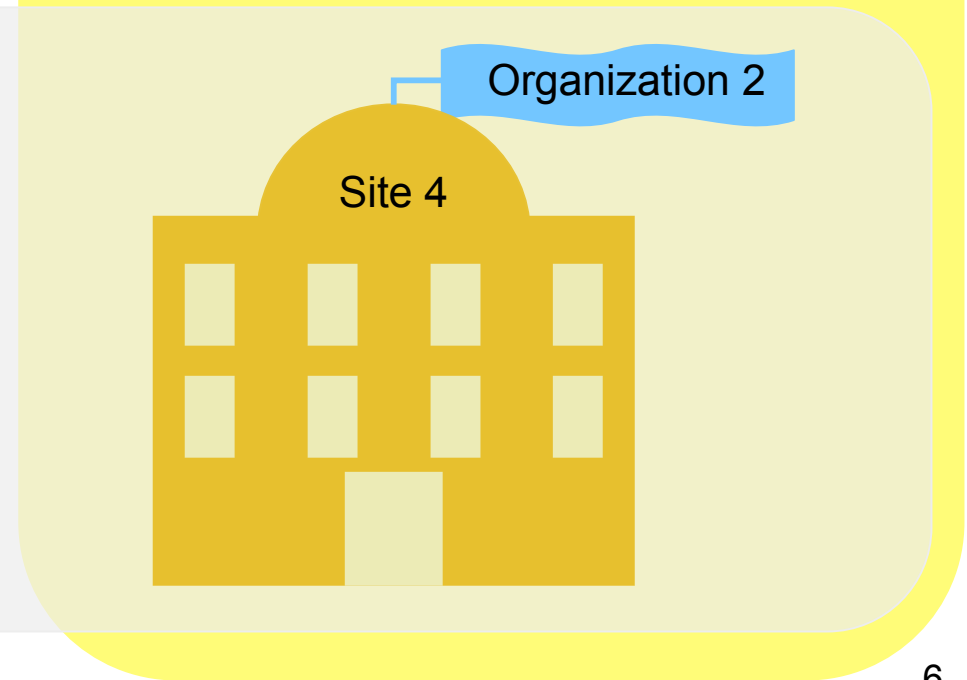
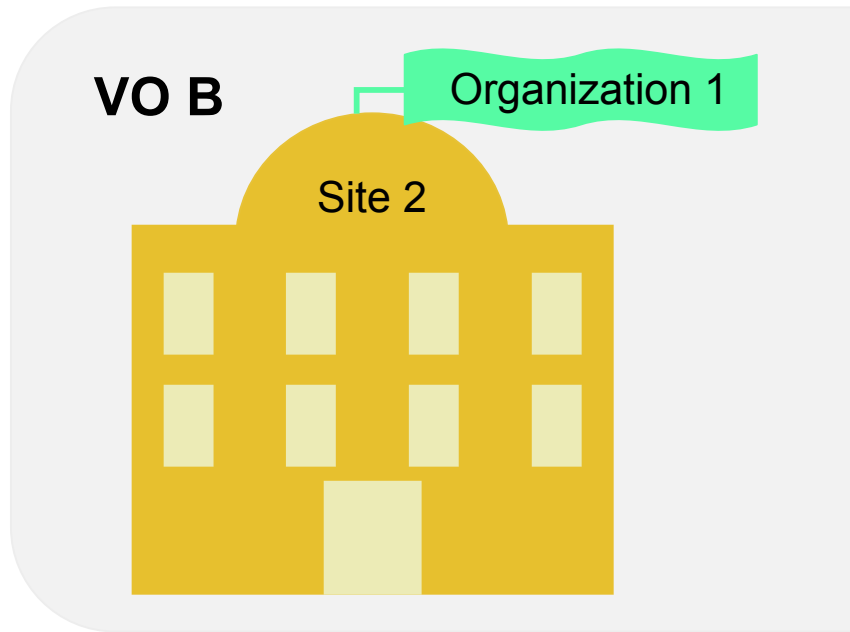
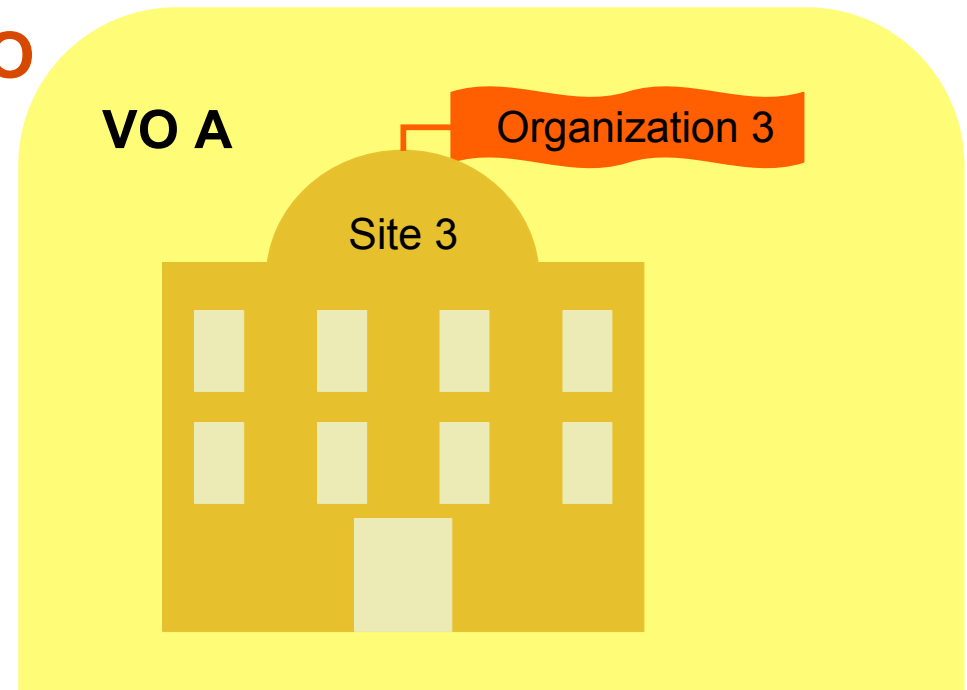
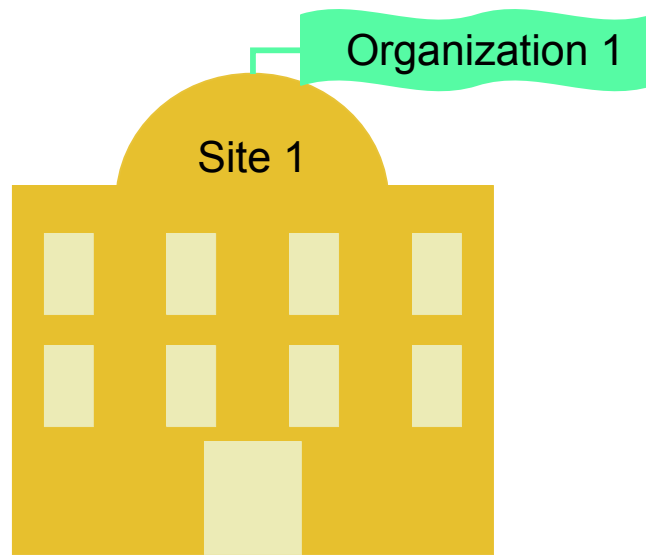


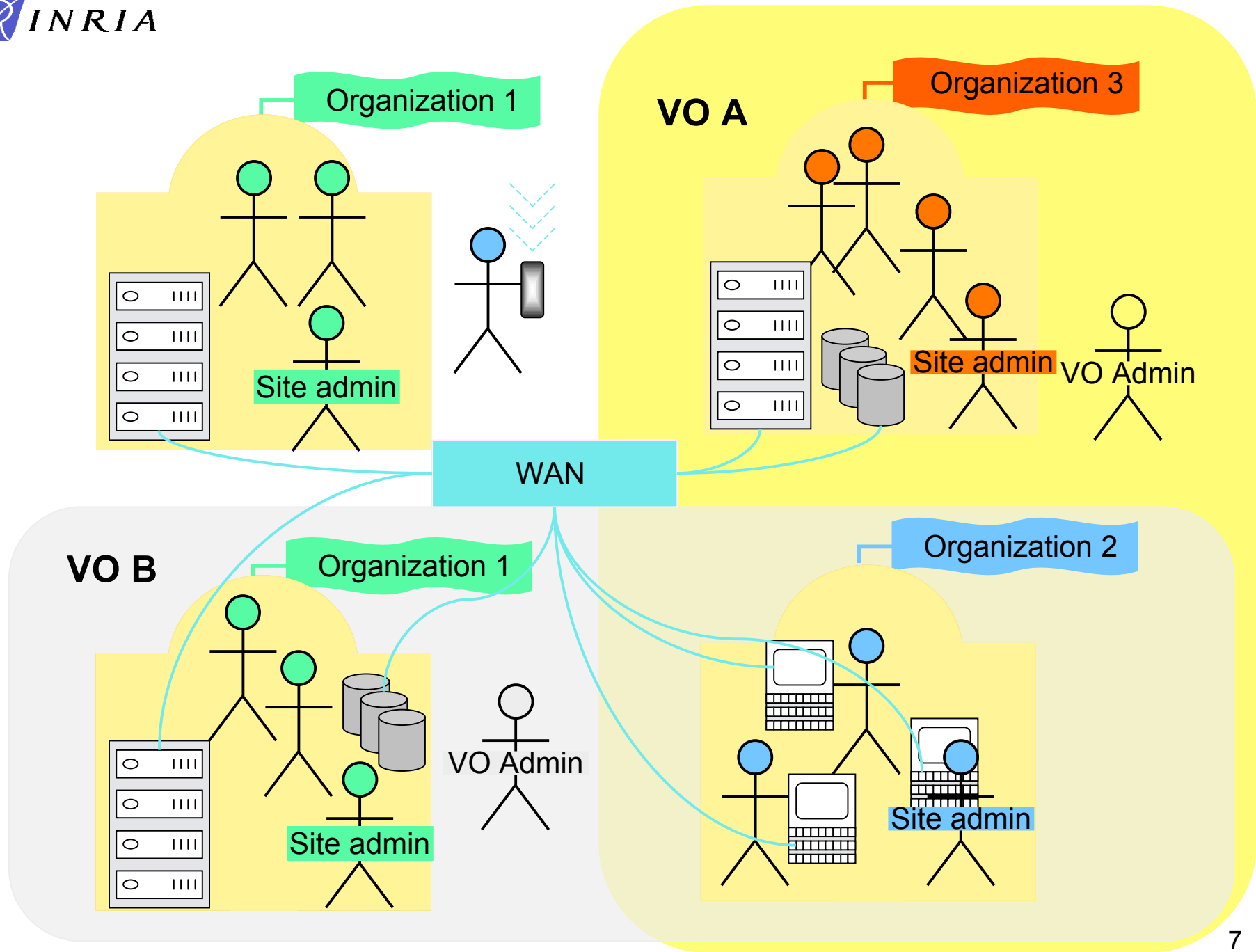
Next Generation Grids



"A fully distributed, dynamically reconfigurable, scalable and autonomous infrastructure to provide location independent, pervasive, reliable, secure and efficient access to a coordinated set of services encapsulating and virtualizing resources (computing power, storage, instruments, data, etc.) in order to generate knowledge"

VO





What are the Actors' Needs?

Users

End users - Service Administrators

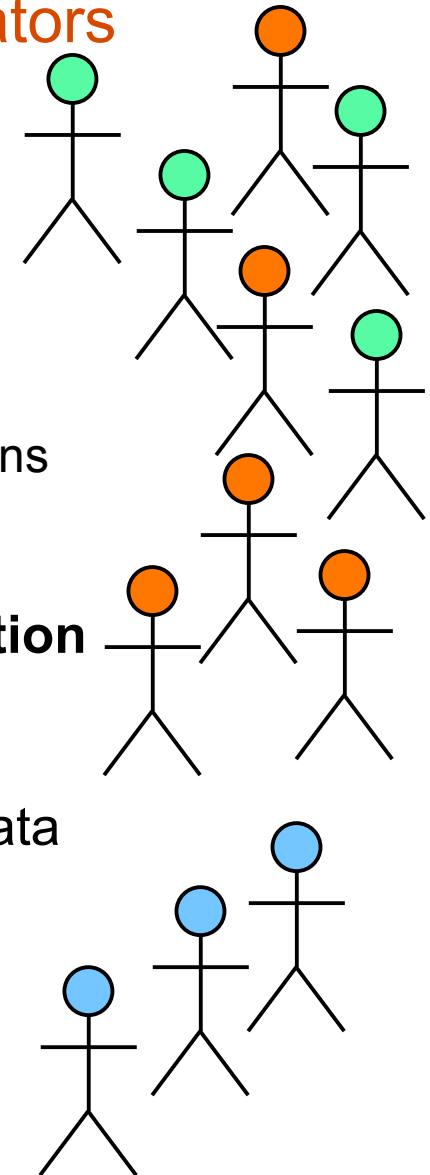
– Ease of use

- Do not want to care with Grid issues
- Want to work with familiar interfaces
- Want to use their non Grid-aware legacy applications
- Simple login as a Grid user in a VO

– Secure and reliable application/service execution

– High performance

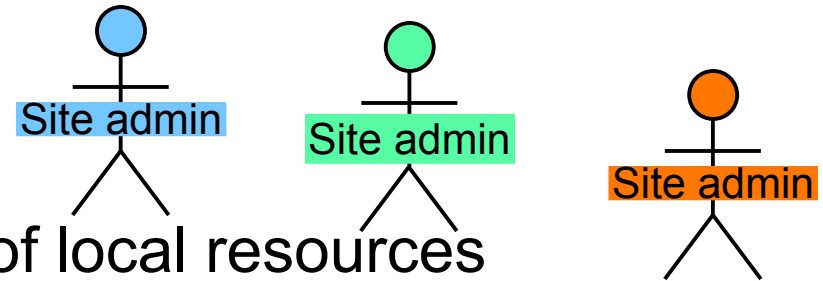
– Ubiquitous access to services, applications & data



Administrators

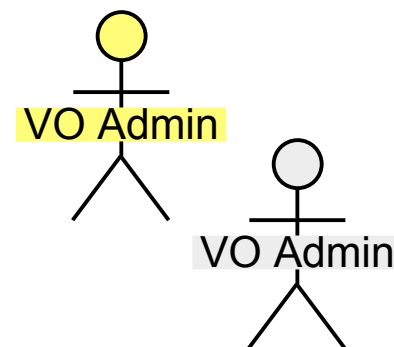
□ Site administrators

- Ease of management
- Autonomous management of local resources
- Should not be impacted by every single change in a VO



□ VO administrators

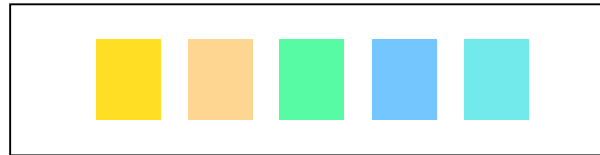
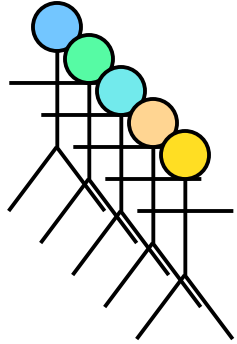
- Ease of management
- Flexibility in VO policies
- Accounting



Developers' Needs

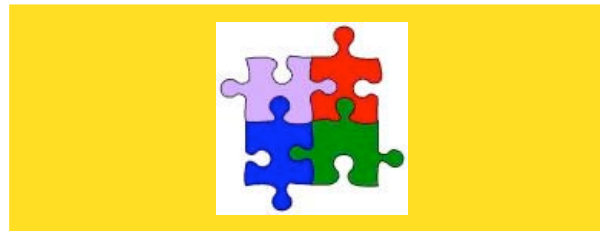
- ❑ **Ease of development of Grid applications**
 - Reuse existing code
- ❑ **Stable API**
- ❑ **Conformance to standard API**
 - Familiar API Posix
 - Grid application standards

Operating System



Application

Set of integrated services
(user account, process, file,
memory segment,
sockets, access rights)



Operating System



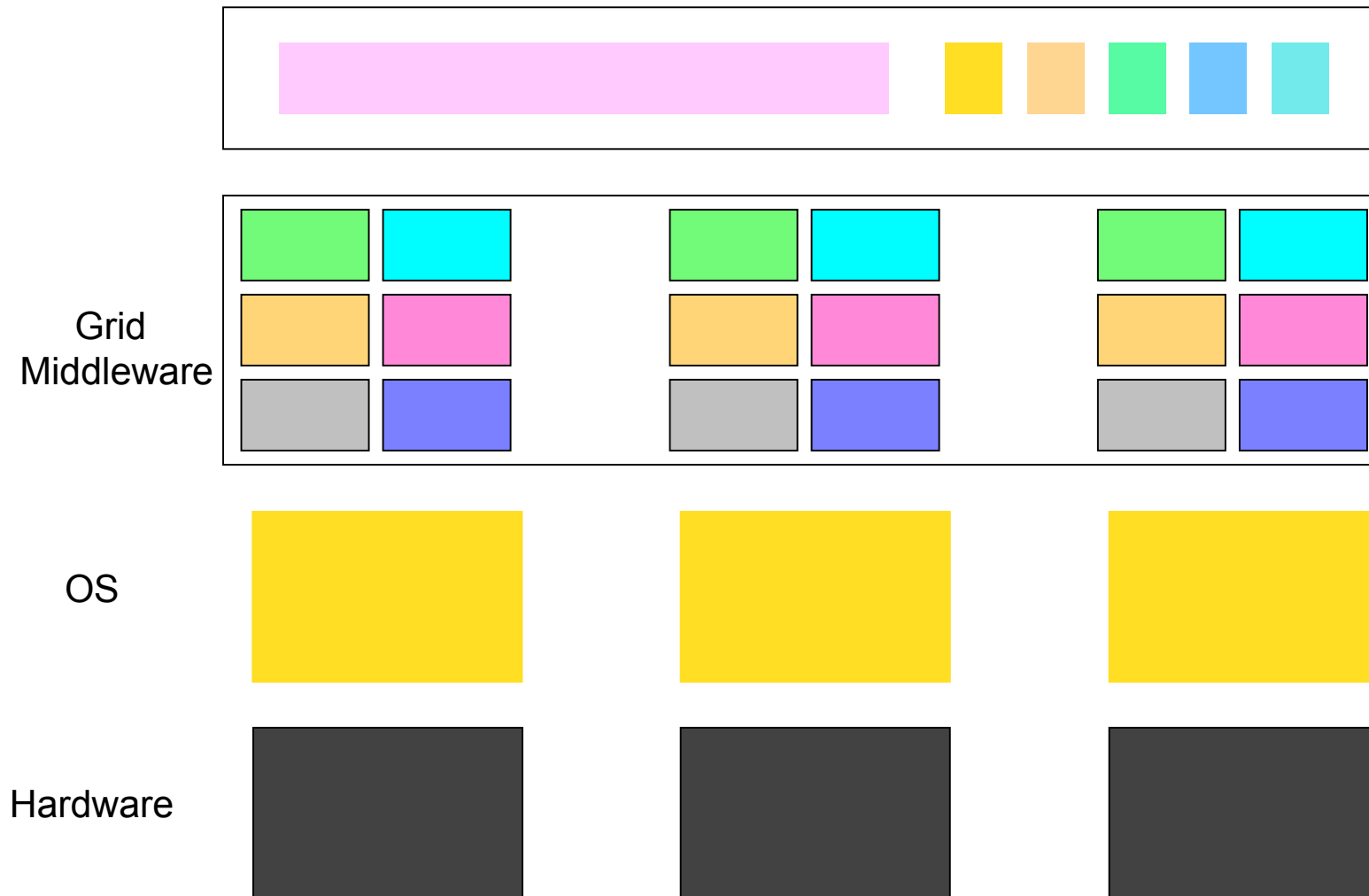
Single computer



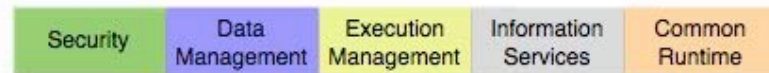
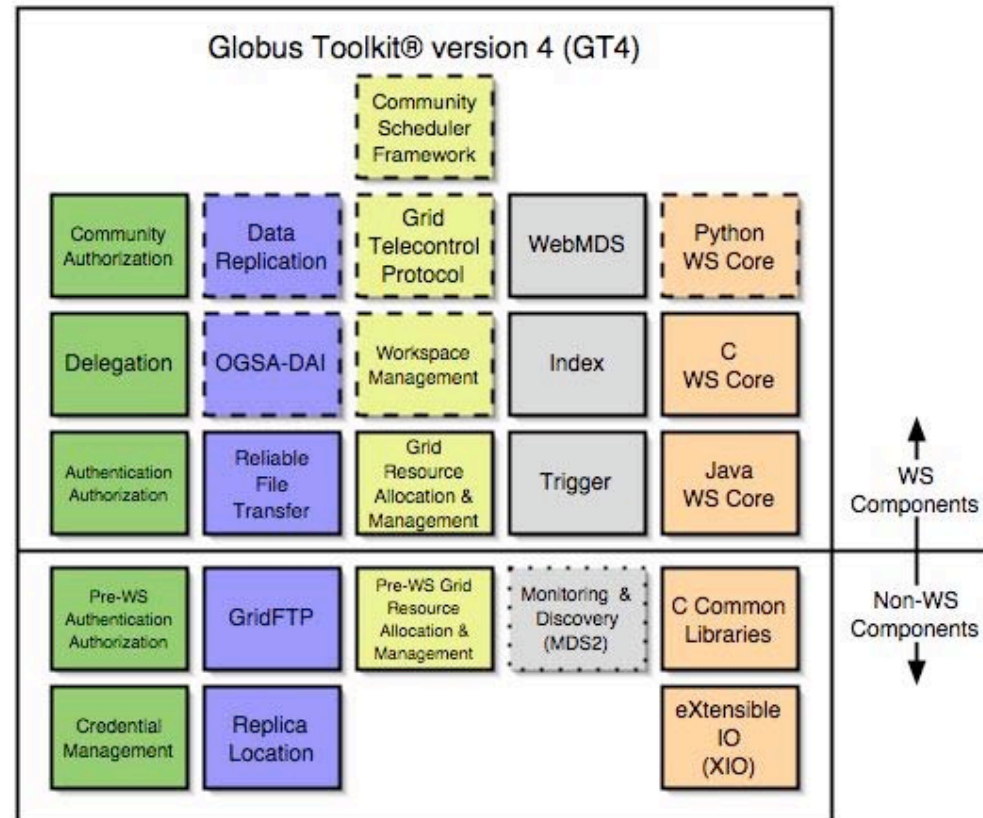
Hardware

Why a Grid Operating System?

Middleware Approach

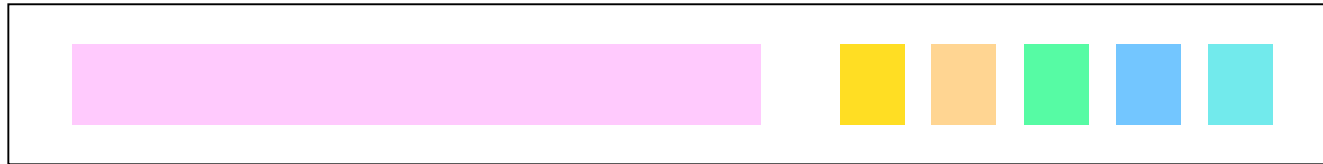


Example: Globus Toolkit

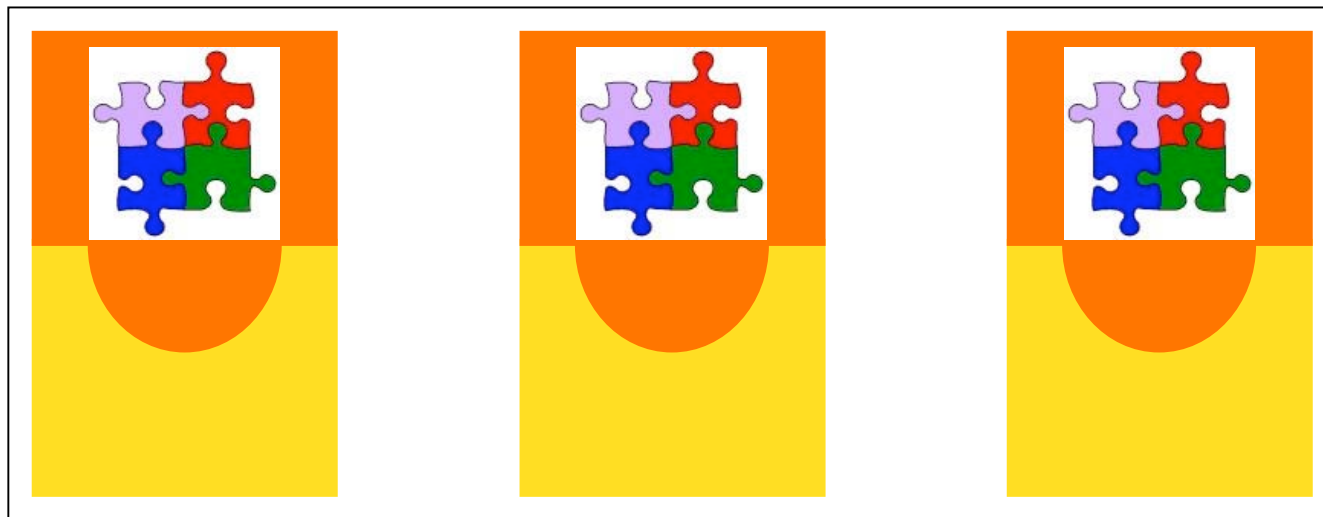


- Core GT Component: public interfaces frozen between incremental releases; best effort support
- Contribution/Tech Preview: public interfaces may change between incremental releases
- Deprecated Component: not supported; will be dropped in a future release

Grid Operating System



Grid OS



Grid Operating System

A **comprehensive** set of **cooperating** system **services**

providing a **stable** interface

for a **wide-area dynamic distributed** infrastructure

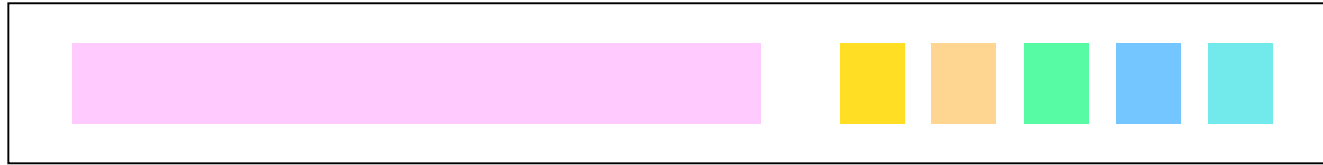
composed of **heterogeneous** resources

spanning **multiple administrative** domains

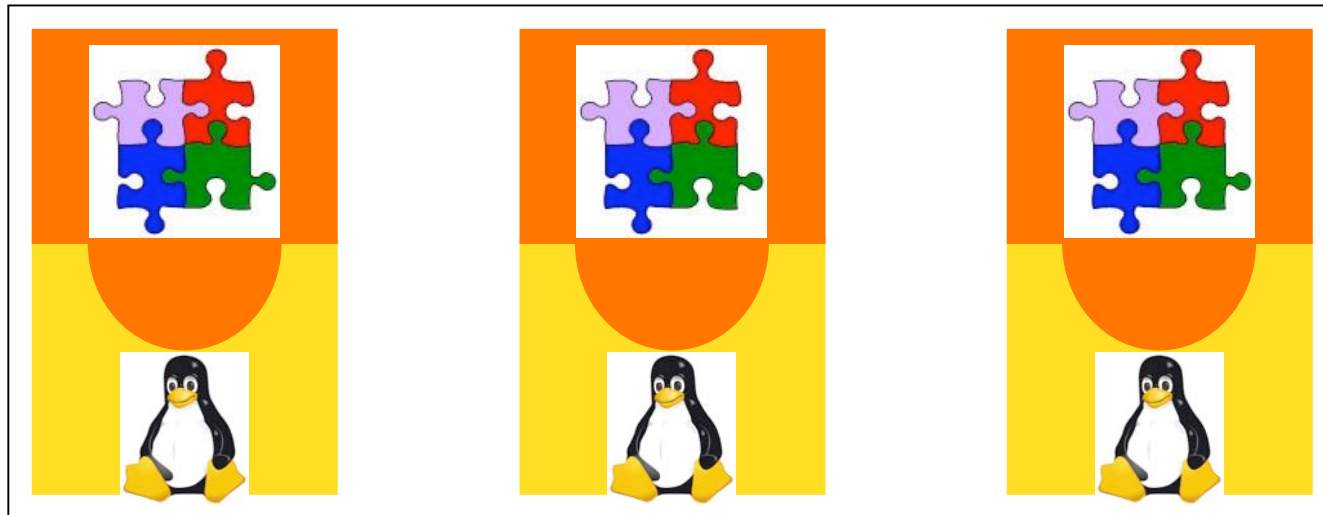
XtreemOS Grid OS

XtreemOS

A Grid OS based on Linux with Native VO Support



XtreemOS



Application Spectrum

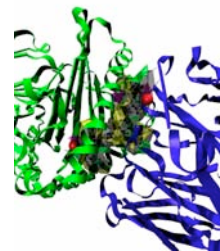
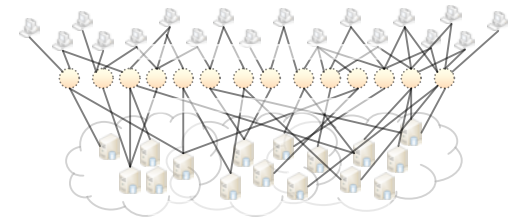
□ Wide range of applications...

- Grid aware distributed applications
- Grid unaware (legacy) applications executed in a Grid

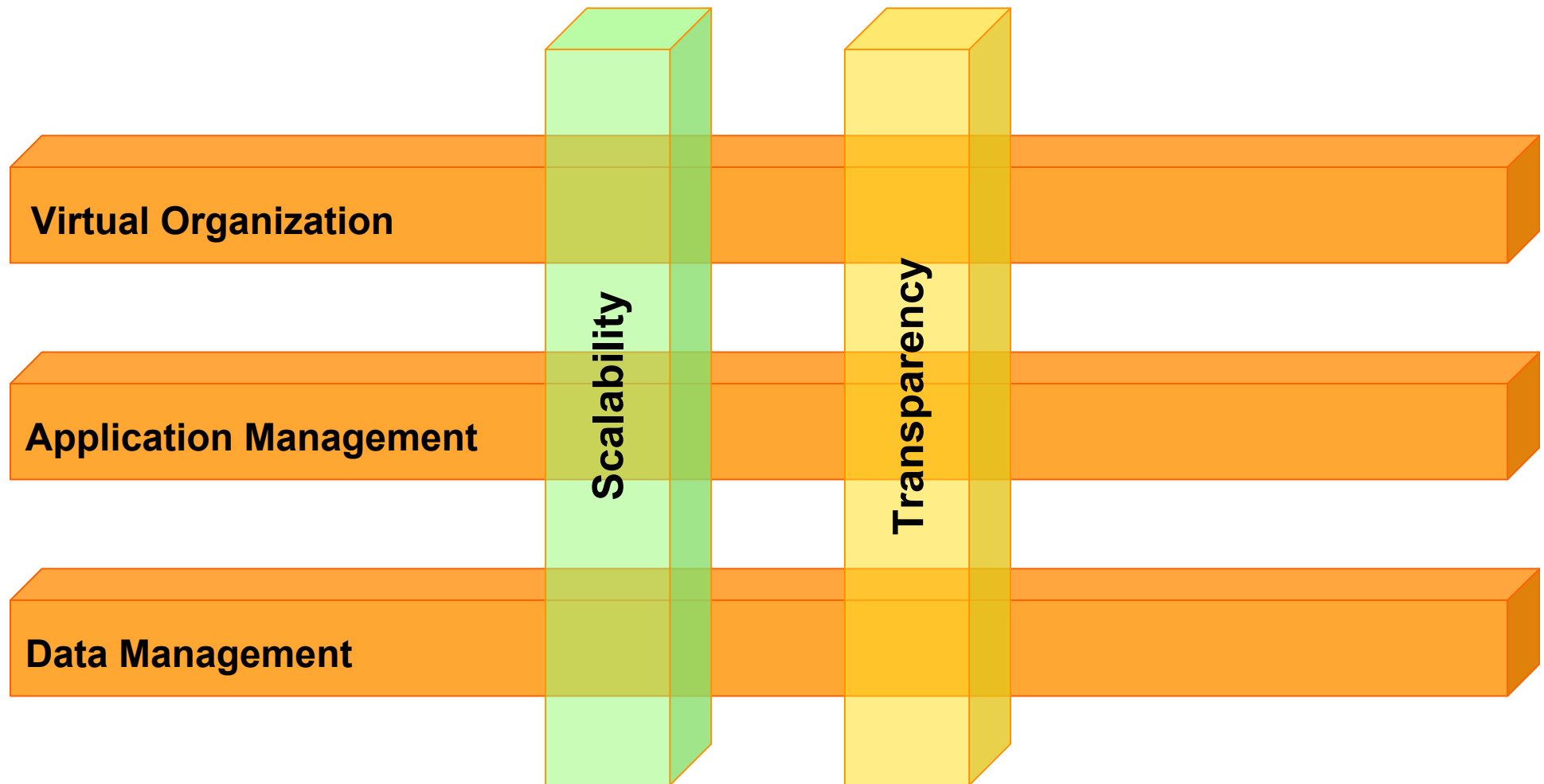
□ ... in different domains

- E-business
 - Services...
- Scientific applications

... **XtreemOS is an OS!**



XtreemOS Fundamental Properties



Scalability

□ Scale

- Thousands of nodes in thousands sites in a wide area infrastructure
- Thousands of users

□ Consequences of scale

- Heterogeneity
 - Node hardware & software configuration
 - Network performance
- Multiple administrative domains
- High churn of nodes

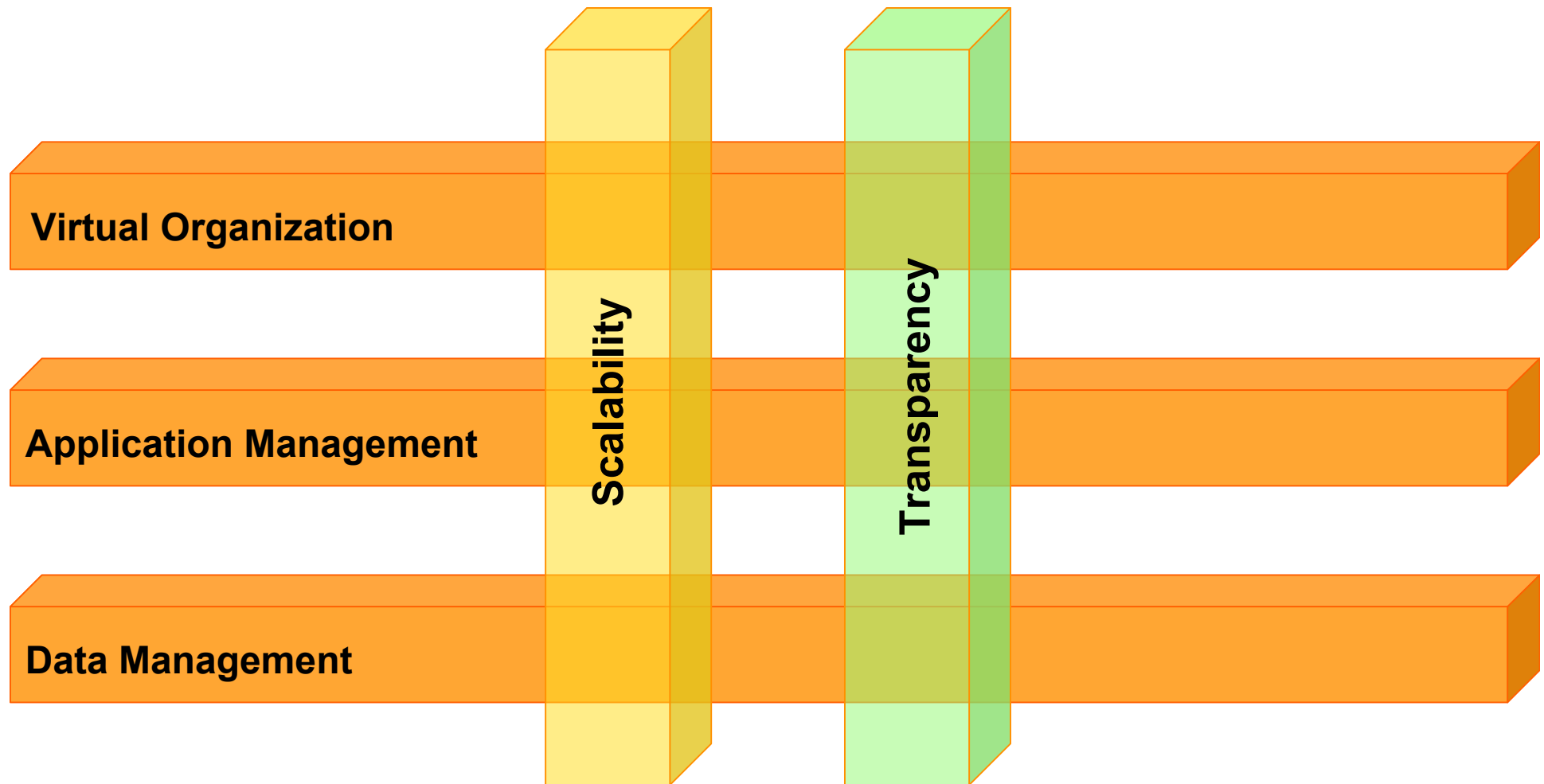


XtreemOS Service Scalability

- ❑ **Scalability with the number of entities & their geographical distribution**
 - Avoid contention points & save network bandwidth (performance)
 - Run over multiple administrative domains (security)

- ❑ **Adaptation to evolving system composition (dynamicity)**
 - Run with partial vision of the system
 - Self-managed services
 - Transparent service migration
 - Critical services highly available
 - No single point of failure

XtreemOS Fundamental Properties



Transparency

User's Point of View

- ❑ **Bring the Grid to standard Linux users**
 - **Feeling to work with a Linux machine**
 - Standard way of launching applications
 - `ps` command to check status of own jobs
 - **No limit on the kind of applications supported**
 - Interactive applications
 - **Grid-aware user sessions**
 - **Grid-aware shell** taking care of Grid related issues
 - **VO can be built to isolate or share resources**
 - Parameter defined by VO administrator

Transparency

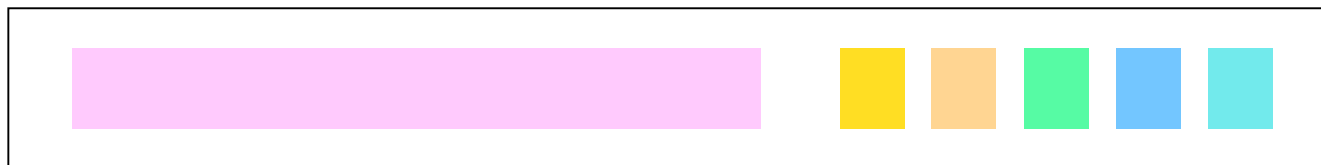
Application & Application Developer's Point of View

- ❑ **Make Grid executions transparent**
 - Hierarchy of jobs in the same way as Unix process hierarchy
 - Same system calls: wait for a job, send signals to a job
 - Processes in a job treated as threads in a Unix process
- ❑ **Files stored in XtremFS Grid file system**
 - Posix interface and semantics to access files regardless of their location
- ❑ **Transparent fault tolerance to applications**
- ❑ **Clusters transparent to applications**
 - Single System Image

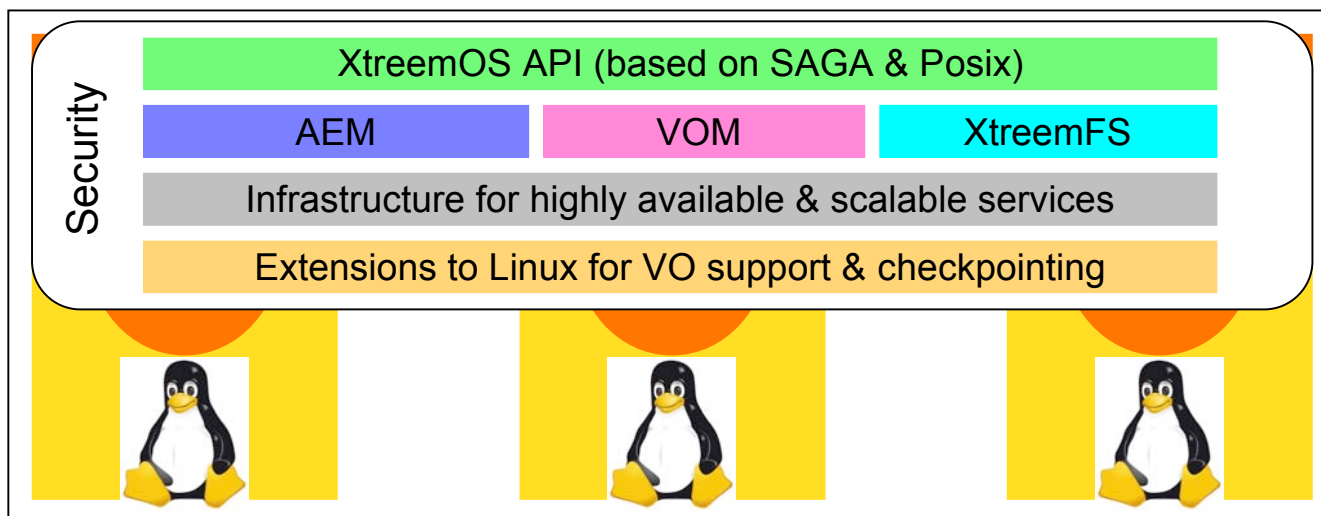
XtreemOS Services

XtreemOS

A VO-aware OS based on Linux



XtreemOS



Virtual Organization Management

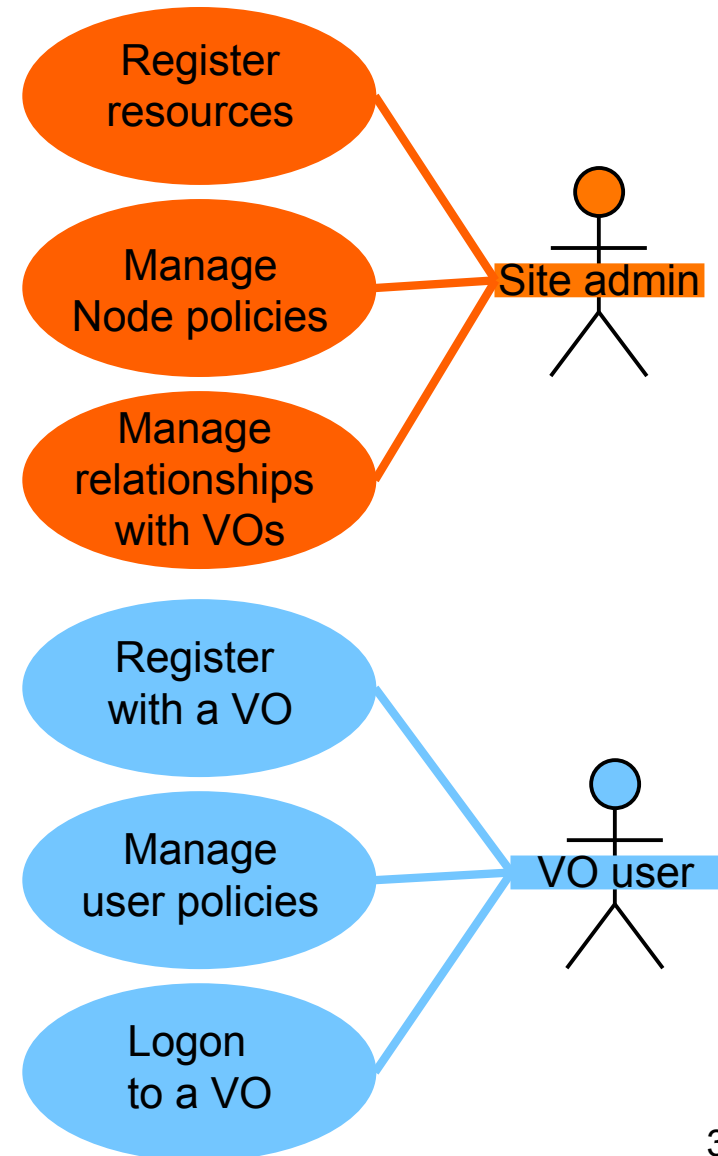
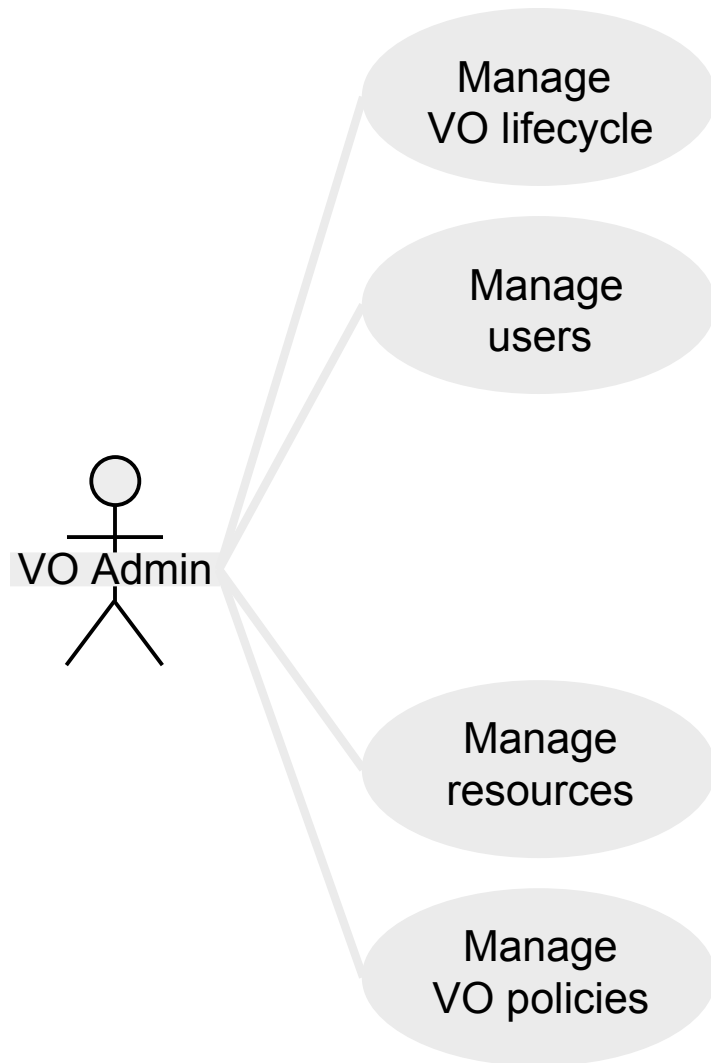
❑ Objectives

- To allow secure interaction between users and resources
 - Authentication, authorization, accounting

❑ Challenges

- Interoperability with diverse VO frameworks and security models
- Flexible administration of VOs
 - Flexibility of policy languages
 - Customizable isolation, access control and auditing
- Scalability of management of dynamic VOs
- Embedded support for VOs in the OS
- No compromise on efficiency, backward compatibility

Use Cases



Security in XtremOS

- ❑ **VO-centric security architecture**
 - **Grid level security services**
 - Global entities: VO, users, nodes (identified by public key certificates)
 - **Node (OS) level services**
 - Local entities: OS users (uid), OS resources (files (inode), process (pid))
 - Hierarchical policy management
 - Resource access control
 - Resource usage
- ❑ **Interoperability with third party security infrastructures**
 - Kerberos, LDAP, Shibboleth...
- ❑ **Single-Sign-On**

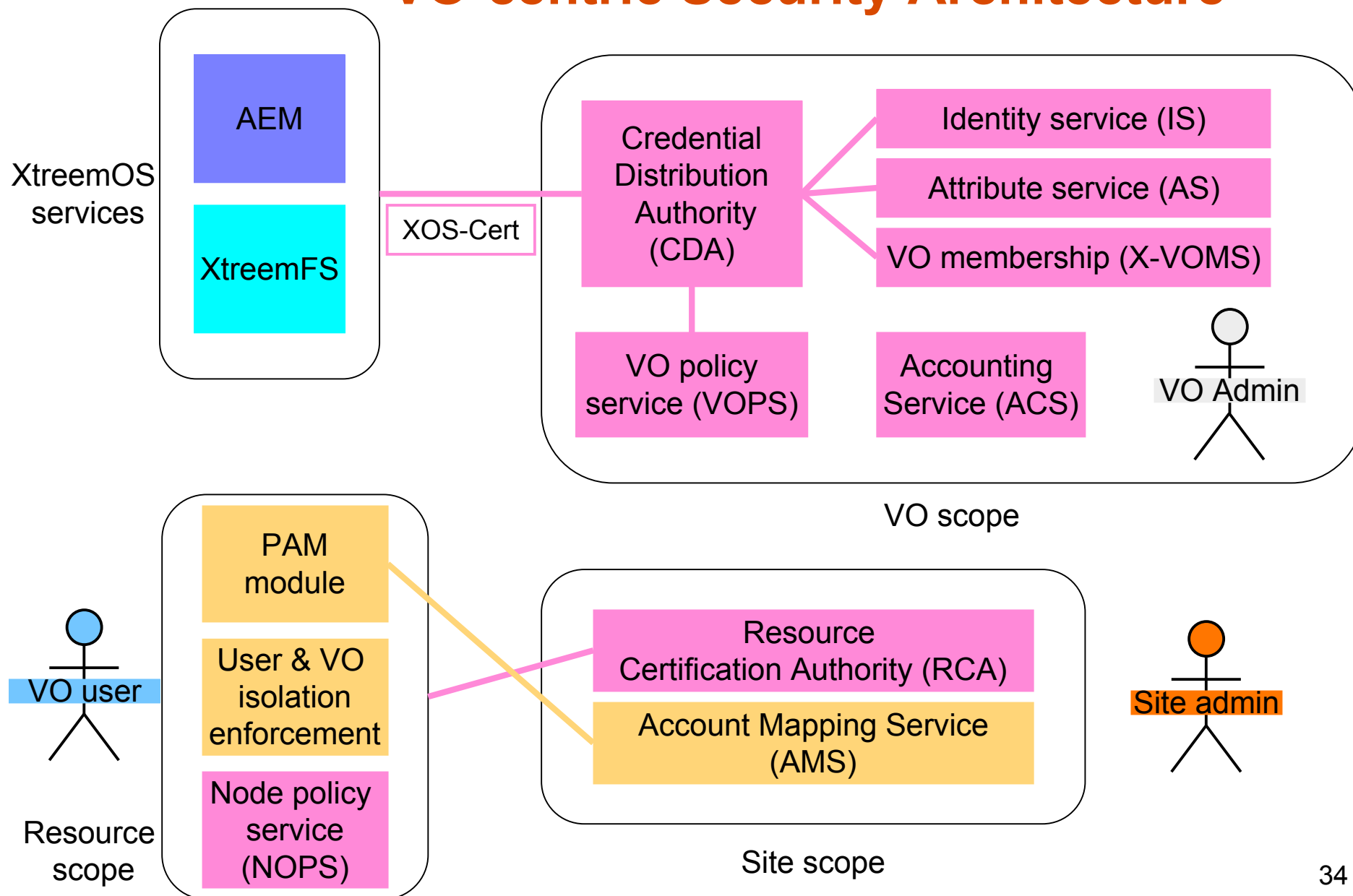
System-Level VO Support

- ❑ **Policies specified by a VO finally checked & ensured at resource nodes by the local instance of the OS**
 - Standard Linux unaware of VOs
 - Isolation & access control mainly rely on user accounts, process id, file permission bits
- ❑ **What is needed for Linux OS to be able to enforce VO policies**
 - OS kernel should deal with VO & VO users identities
 - Identity information should be exploited in standard access control mechanisms
 - Linux OS should supply identity information to Grid level services (XtreemFS, AEM)
- ❑ **NO modification of Linux kernel**
 - Mapping of VO level identities & policies into local ones fully recognized by Linux

System-Level VO Support

- ❑ **VO-customizable, dynamic mapping of Grid users onto local accounts**
 - Integration of Grid user management into Linux using
 - Pluggable Authentication Modules (PAM)
 - Multiple low level authentication technologies into a common high level API
 - Name Service Switch (NSS)
- ❑ **Interfacing with the Grid authentication services**
 - Development of PAM modules to accommodate multiple VO models
 - Authentication, authorization, session management
- ❑ **User space credential translation**
 - NS-Switch
- ❑ **Access control & logging**
 - Caching of authentication data related to a process within the kernel

VO-centric Security Architecture

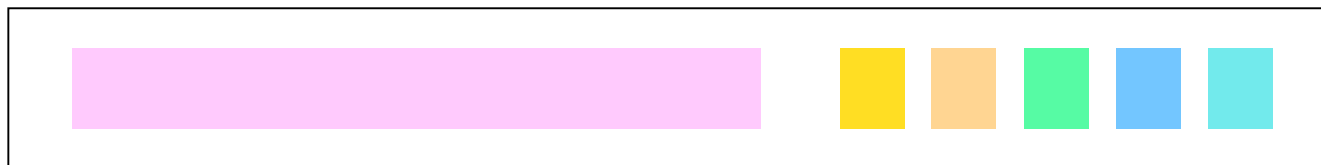


Key Contributions

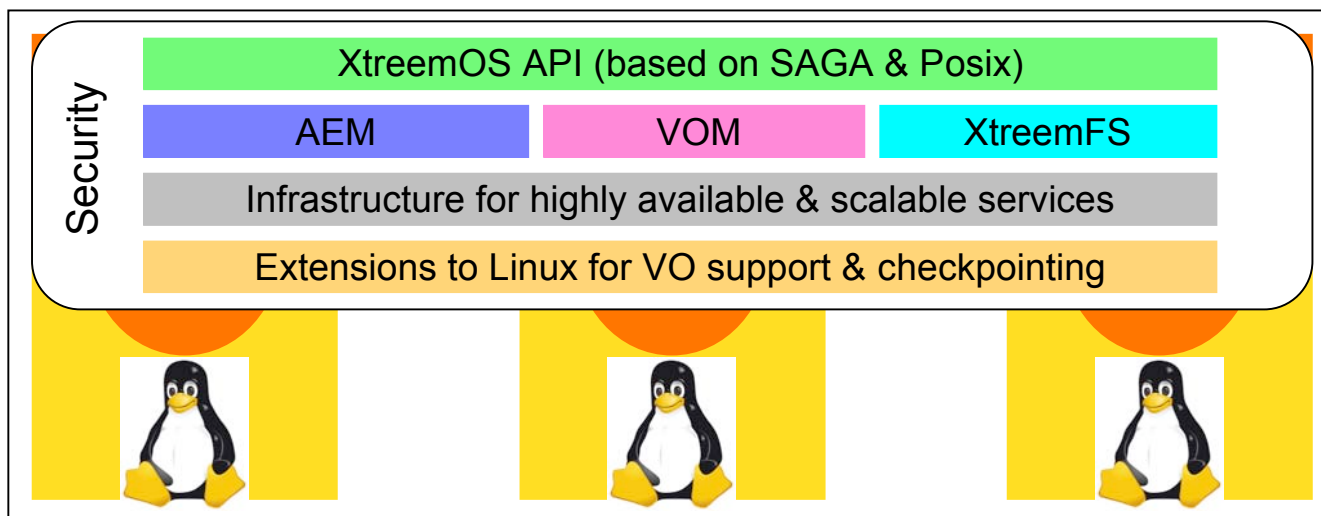
- ❑ **Maximum transparency**
 - Grid unaware applications & tools can be used without being modified or recompiled
- ❑ **Integration of Grid level authentication with system level authentication**
 - Creation of dynamic on-the-fly mappings for Grid users in a clean & scalable way
 - No centralized Grid wide data base
- ❑ **Grid user mappings invisible to local users**
- ❑ **VO are easier to setup and manage**
 - No grid map file needed
 - User management does not necessitate any resource reconfiguration

XtreemOS

Application Execution Management



XtreemOS



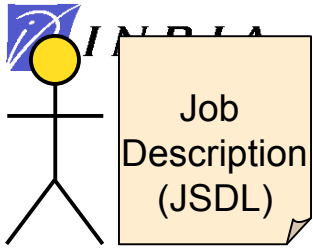
Application Execution Management

□ Objectives

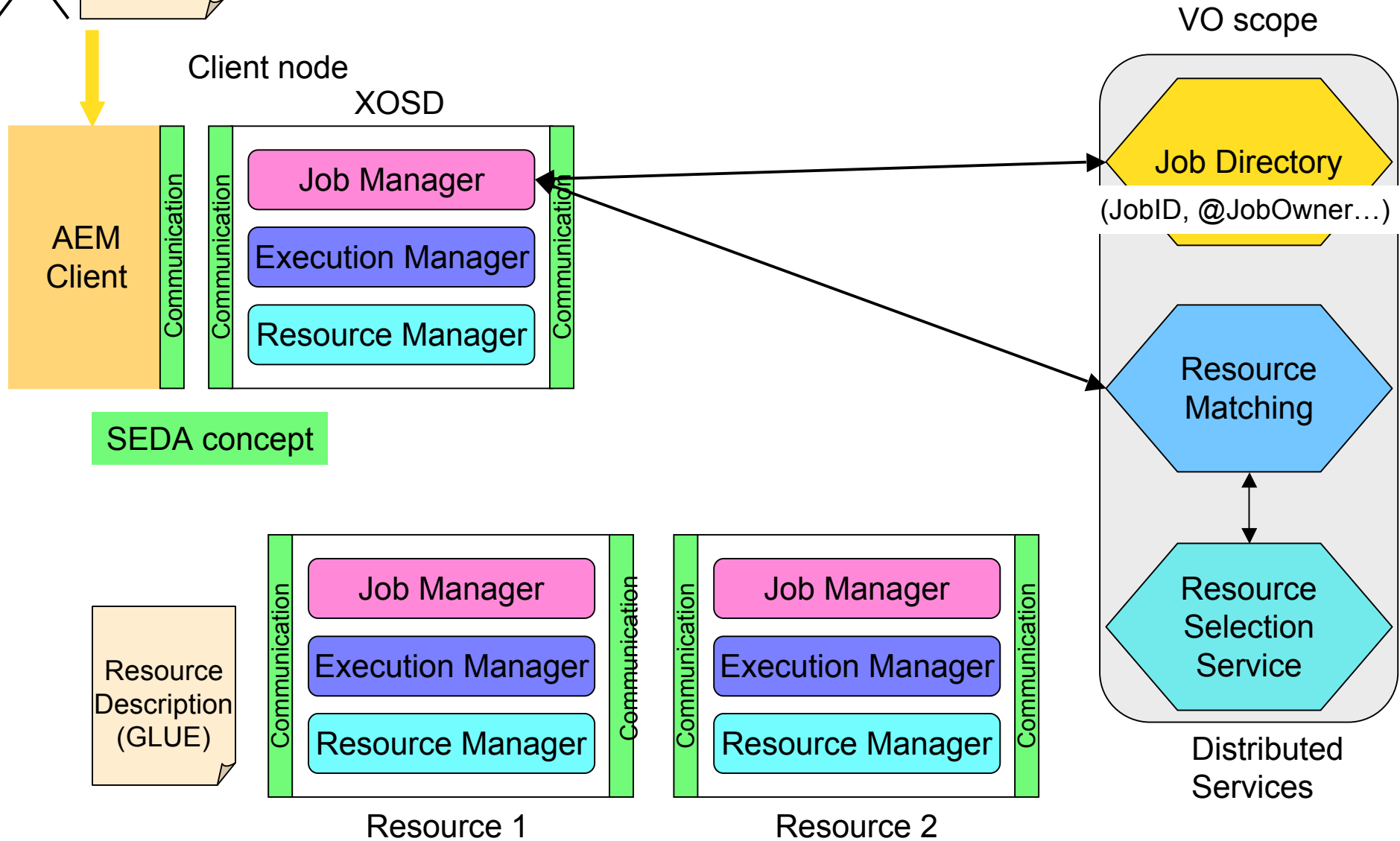
- Start, monitor, control applications
- Discover, select, allocate resources to applications

□ Challenges

- Deal with a large variety of resources with changing conditions over time
- Cost to obtain system information and take appropriate decisions has to be orders of magnitude less than in Grid middleware-based systems
- Take advantage of accurate information for better scheduling control



AEM Architecture



Main Features

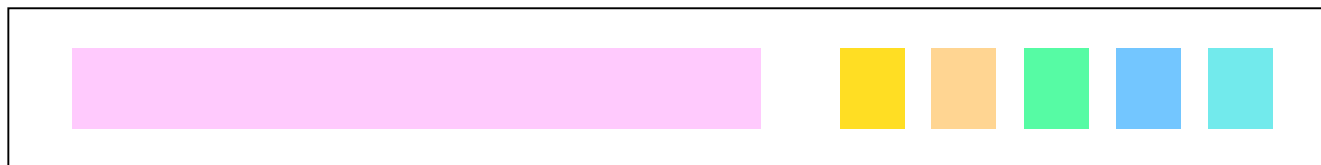
- ❑ **No global job scheduler**
- ❑ **Distributed management of jobs**
- ❑ **No assumption on local node RMS**
 - AEM can be used without any batch system
- ❑ **Resource discovery based on overlay networks**
 - Structured and unstructured
 - Multi-criteria and range of values queries

Advanced Features

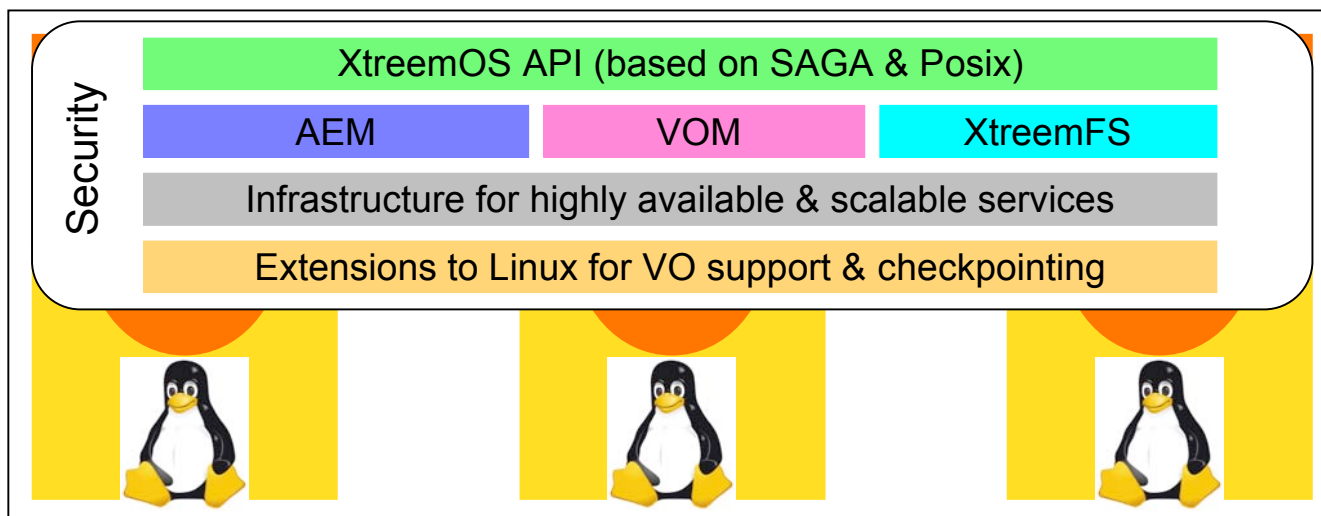
- ❑ **Flexible monitoring**
- ❑ **Accounting**
- ❑ **Reservation**
 - Nodes with a local resource manager
 - Co-allocation of resources
- ❑ **Checkpoint/restart mechanisms for grid jobs**
- ❑ **Migration of grid jobs when the user agreement cannot be met anymore**
- ❑ **Interactive applications support**
- ❑ **Support for external workflow engines**

XtreemOS

Data Management



XtreemOS



Data Management

❑ Objectives

- Providing to users a global view of their files & transparent access to data through a Grid file system

❑ Challenges

- Efficient location-independent access to data through standard Posix interface in a Grid environment
 - Grid users from multiple VO
 - Data storage in different administrative domains
- Autonomous data management with self-organized replication and distribution
- Consistent data sharing
- Advanced meta data management

Conclusion

- ❑ **XtremOS is not yet another Grid middleware**
 - **Operating system** for large scale wide-area platforms distributed over multiple administrative domains
 - Comprehensive set of cooperating services
 - Stable Posix interface
 - **Grid-aware Linux distribution**

- ❑ **Native Virtual Organization Support**
 - Flexible & scalable VO management
 - Multi-VO & short-term VO support

- ❑ **Secure, reliable, efficient application/service execution & ease of use and management**

- ❑ **Attractive in the context of the new emerging computing models**



Get Involved!

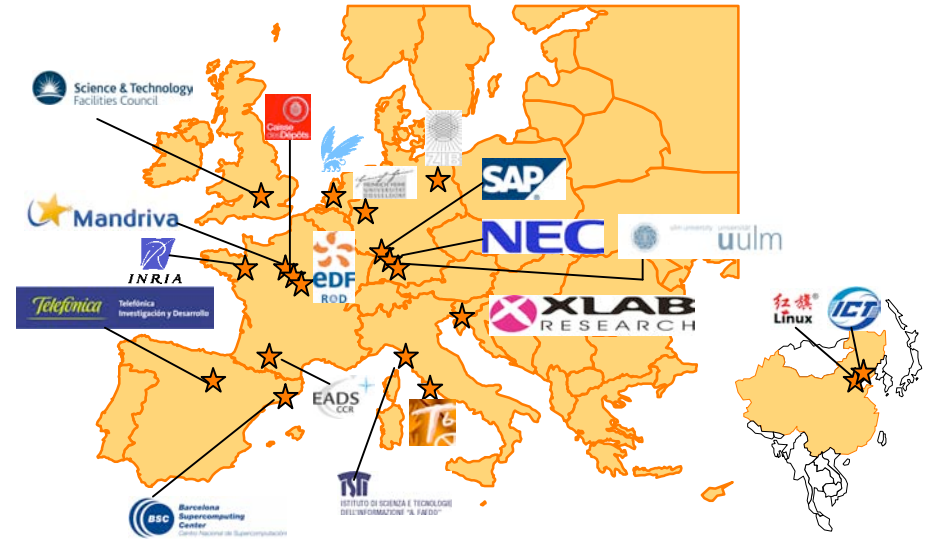
- ❑ **Download the first XtreemOS public release in a few days (GPL/BSD)**
 - <http://www.xtreemos.eu>
 - **Open development**

- ❑ **`contact@xtreemos.eu` to register in the pioneer user group**



Acknowledgements

- XtreamOS consortium
 - <http://www.xtreemos.eu>



- PARIS project-team @ INRIA Rennes - Bretagne Atlantique
 - <http://www.irisa.fr/paris>





Thank you for your Attention

<http://www.xtreemos.eu>