Grid Computing

Grid Computing

- Definitionen
- Globus Dienste
- Verwendung
- Programmierung

Definitionen

„A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high-end computational capabilities“

I. Foster (1999)

Definitionen

GRID COMPUTING:

„The real an specific problem that underlies the grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organisations. The sharing we are concerned with is not primarily file exchange but rather dierct access to computers, software, data, and other resources as is required by a range of collaborative problemsolving and resource-brokering strategies emerging in industry, science and engineering. The sharing is is necessarily, highly controlled, with resource providers and consumers defininig clearly and carefully just what is shared, who is allowed to share and the conditions under which sharing occurs. A set of individuals and/or institutions defined by such sharing rules form what we call a virtual organisation.“

I. Foster (2001)

Definitionen

- A grid is a system that ...
  - ... coordinates resources that are not subject to centralized control.
  - ... uses standard, open, general-purpose protocols and interfacecs.
  - ... delivers nontrivial qualities of service

(I. Foster 2002)

Definitionen

- Virtual Organization:
  - Menge von Personen und oder Organisationen
  - Teilen sich Ressourcen
  - Auf kontrollierte Weise
  - Zusammenarbeit zum Zwecke der Erfüllung einer berechnungs- und/oder datenintensiven Aufgabe
  - Temporär

Definitionen

- Ressource:
  - Computer
  - Netzwerke
  - Datenarchive
  - Wissenschaftliche Geräte
  - Visualisierungsgeräte
  - ...
- Je nach Betrachtungsebene
  - Physikalische Einheit
  - Logische Einheit

„Grid“

- Begriff Grid in Analogie zum „electrical power grid“
- Grid als Infrastruktur:
  - Verlässlichkeit (Sicherheit, Verfügbarkeit)
  - Konsistenz (Standardisierte Schnittstellen)
  - Durchdringung („Verfügbarkeit überall“)
    - Vergleichbar mit Stromnetz
  - Günstiger Zugang

Eigenschaften und Probleme von Grids

- Größe und Notwendigkeit der (Ressource)selektion
- Heterogenität (Hardware, Betriebssystem, Scheduler, Verwendungsregelungen)
- Dynamisches und unvorhersehbares Verhalten (Teilen von Ressourcen)
- Mehrere administrative Bereiche

Notwendige Technologien für VOs und Grids

- Sicherheitslösungen
- Management von Berechtigungen und Regelungen über mehrere Organisationen hinweg
- Ressource Management Protokolle und Dienste, die den sicheren entfernten Zugriff zu Berechnungs- und Datenressourcen sowie die gleichzeitige Allokation von mehreren Ressourcen unterstützen
- Informationsabfrageprotokolle und Dienste, die Status- und Konfigurationsinformationen zu Ressourcen, Organisationen und Services bieten
- Datenmanagementservices die Daten finden und zwischen Speichersystemen und Anwendungen transportieren.

Globus Toolkit

- Grundlegende Software-Infrastruktur für Konstruktion und Einsatz von portablen und (effizienten) Diensten
- Menge von Low-Level Mechanismen als Bausteine
- Möglichkeiten zur Beobachtung und Steuerung der Low-Level Mechanismen

Globus Toolkit

- Wichtige Globus Services (GT2):
  - Globus Resource Allocation Manager (GRAM)
  - Globus Security Infrastructure (GSI)
  - Globus Metadirectory Service (MDS)
  - GridFTP, Globus GASS, Globus IO

GRAM

- Ausführen von Jobs auf entfernten Rechnern
- Schnittstelle zwischen lokalem Ressource Management System und Grid
- Involvierte Komponenten:
  - Globus Gatekeeper
  - Globus Security Infrastructure
  - Globus Jobmanager

Grundsätzliche Funktionsweise

1. GRAM Client verbindet sich zu remote Gatekeeper
2. Wechselseitige Authentifizierung (GSI)
3. Übertragung einer Anforderung in der Resource Specification Language (RSL)
4. Gatekeeper bestimmt über GSI lokalen Benutzernamen (nach Zertifikat)
5. Jobmanager wird gestartet und bekommt Job zugewiesen
6. Jobmanager führt Job unter lokalem Benutzer aus

GRAM

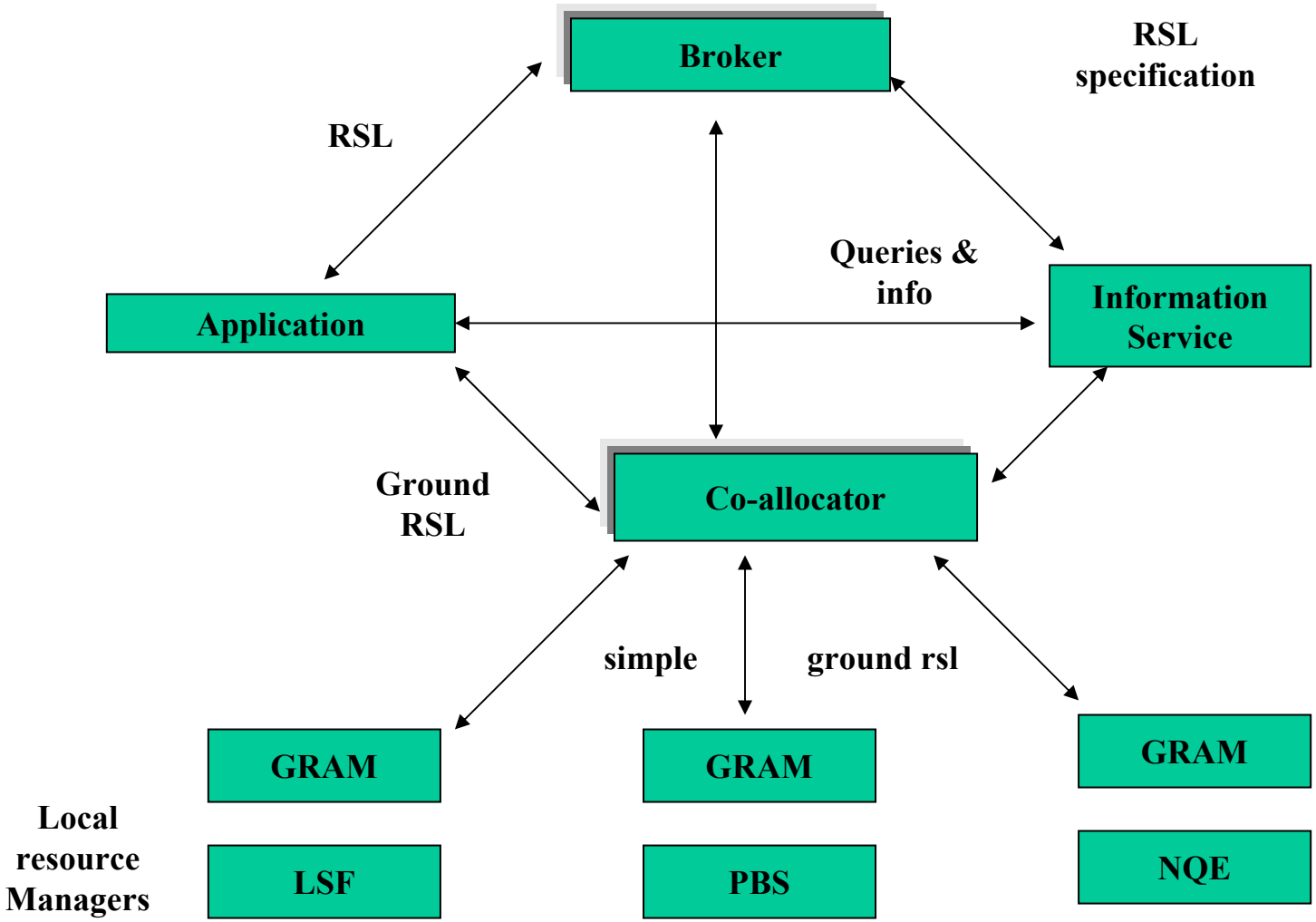
- GRAM Repoter liefert aktuelle Resourceinformationen an MDS:
  - Systemlast
  - Zahl der Jobs
  - Auch statische Resourceinfos
  - Reservierungsunterstützung durch Scheduler
- Infos für Resource Brokers
  - Nicht Globus Core Service
    - Nimrod-G
    - AppLeS
    - Condor DAGman

GRAM

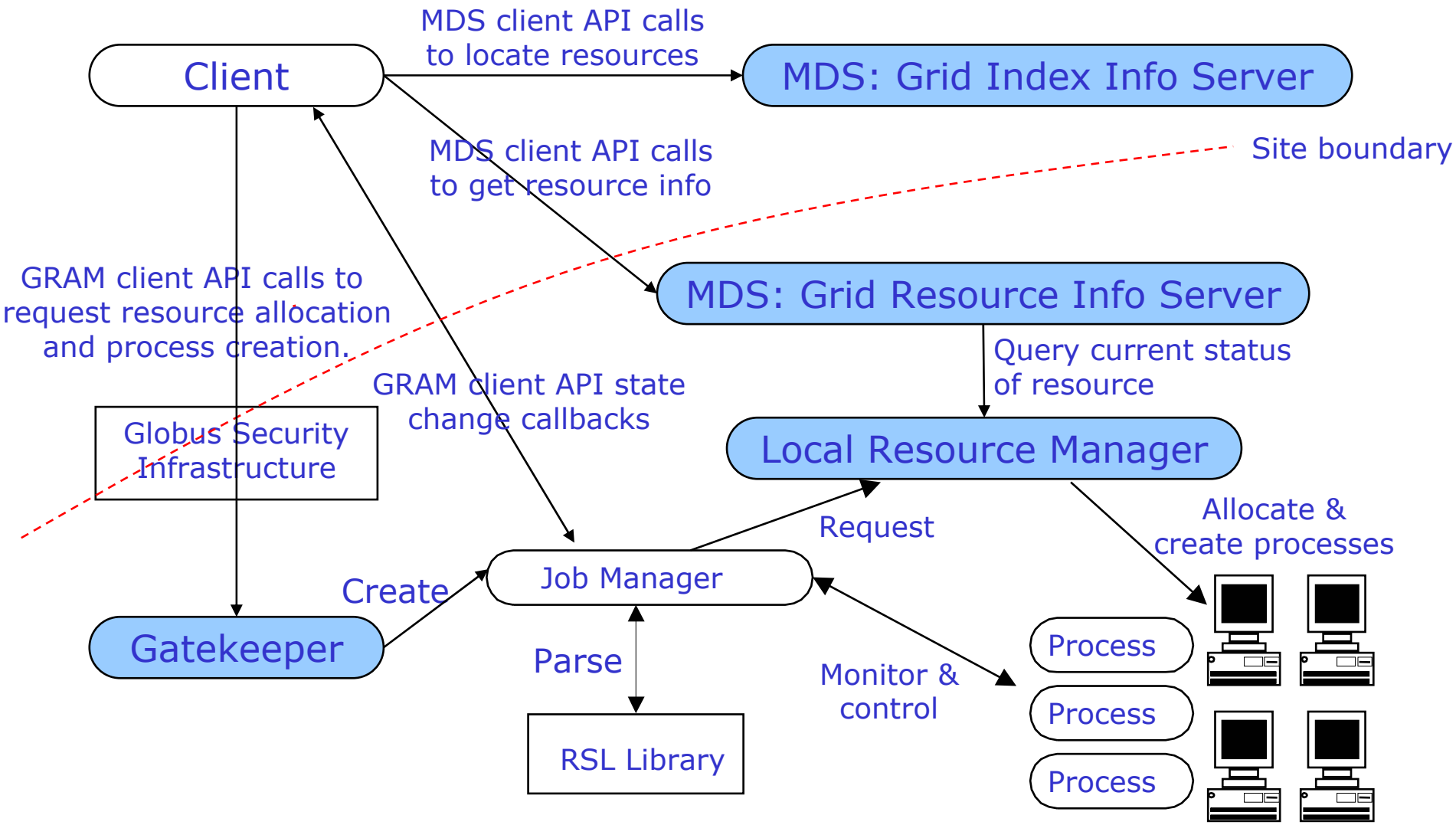
Resource Co-Allocation:

- Ressource Broker oder Benutzer:
  Verwendung mehrerer Resourcen
  (z.B. MPI)
- Resource Co-Allocator zerlegt Anforderung nach Resourcen
- Resource Co-Allocator ermöglicht gemeinsames Monitoring
- Gleichzeitige Resource Allokation
- Lokaler Fehler => Globaler Fehler

GRAM


Broker
RSL specification
RSL
Queries & info
Application
Information Service
Ground RSL
Co-allocator
simple
ground rsl
GRAM
GRAM
GRAM
Local resource Managers
LSF
PBS
NQE

GRAM


Client
MDS client API calls to locate resources
MDS: Grid Index Info Server
MDS client API calls to get resource info
Site boundary
GRAM client API calls to request resource allocation and process creation.
MDS: Grid Resource Info Server
Query current status of resource
GRAM client API state change callbacks
Globus Security Infrastructure
Local Resource Manager
Request
Allocate & create processes
Create
Job Manager
Gatekeeper
Parse
Monitor & control
Process
Process
Process
RSL Library

GSI

- Sicherheit:
  - Authentifizierung
  - Authorisierung
  - Verschlüsselung
- Globus:
  - Authentifizierung (Ressourcen & Benutzer)
  - Verschlüsselung
- Lokal:
  - Authorisierung

GSI

- Probleme im Grid:
  - Große und dynamische Benutzermenge
  - Großer und dynamischer Ressourcepool
  - Dynamische, verteilte Prozessgruppen
  - Heterogene lokale Authentifizierung und Authorisierung
  - Umgebung für einen Benutzer pro Ressource anders (userid, ...)

GSI

- Anforderungen:
  - Einmalige Anmeldung
  - Schutz von Passwörtern, privaten Schlüsseln
  - Interoperabilität
  - Einheitliche Zertifizierungsinfrastruktur

GSI

- Zertifizierungsinfrastruktur:
  - Public Key Infrastructure
  - Certification Authorities (CAs)
    - Garantierte Zuordnung Benutzer ⇔ Zertifikat
  - Asymetrische Kryptographie
  - Vergeben X.509 Zertifikate (Public Key)
  - Identifikation von Ressourcen und Benutzern

GSI

- Lokale Kontrolle (Ressource):
  - CA des Benutzerzertifikates muss von der Ressource akzeptiert werden
  - Abbildung auf lokalen Benutzer
- Benutzer muss die CA der Resource akzeptieren

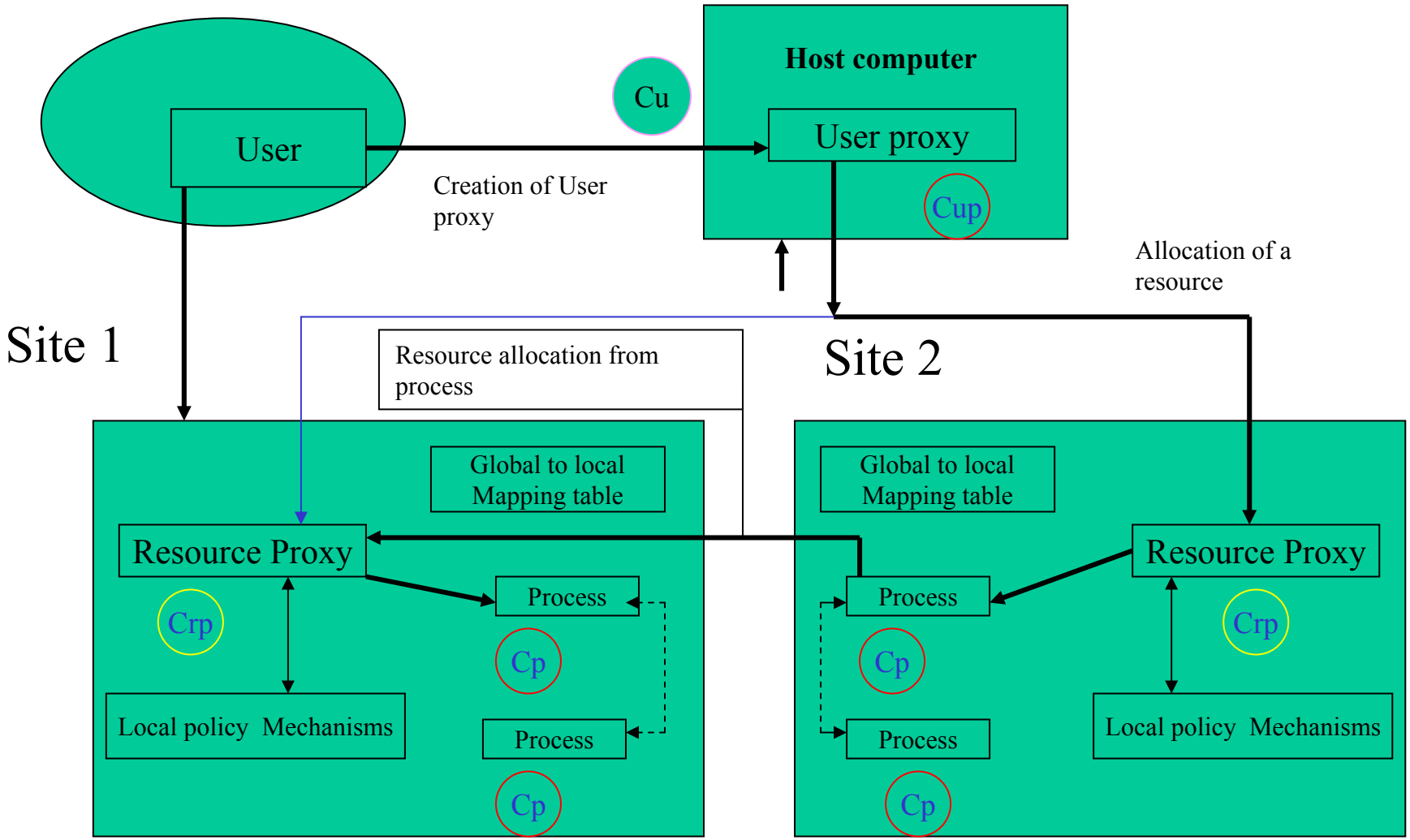
  => wechselseitige Akzeptanz

  => Authentifizierung

GSI

- Authorisierung:
  - Abbildung des Grid-Benutzers auf einen lokalen Benutzer
  - Zugriffsrechte werden lokal auf der Ressource verwaltet
- Verschlüsselung:
  - Datenverschlüsselung für Transfer
  - SSL / TLS

GSI


User
Cu
Host computer
User proxy
Cup
Creation of User proxy
Allocation of a resource
Site 1
Site 2
Resource allocation from process
Global to local Mapping table
Resource Proxy
Crp
Process
Cp
Local policy Mechanisms
Process
Cp
Global to local Mapping table
Process
Cp
Process
Cp
Resource Proxy
Crp
Local policy Mechanisms

MDS

- Liefert Informationen zu Ressourcen in einer VO
- Hierarchisches System
  - Grid Ressource Information Service (GRIS), Ressource-bezogen
  - Grid Index Information Services (GIIS) VO-weite Informationen
- (Benutzer)-Zugriff über LDAP

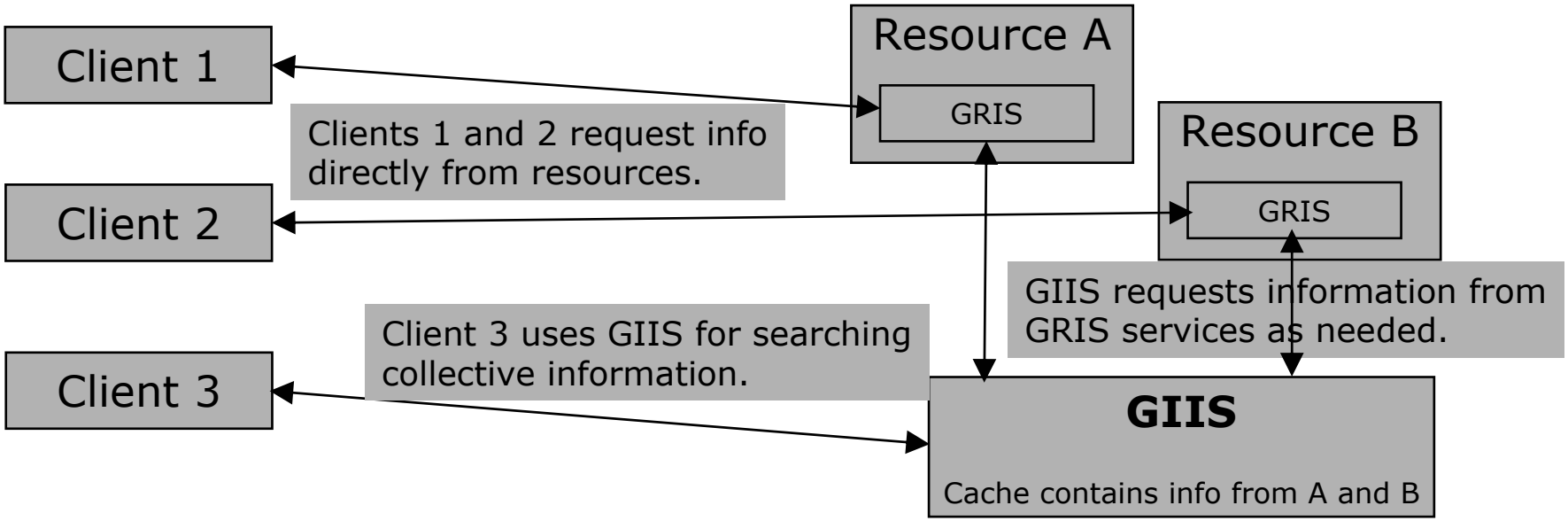
MDS

- Information Providers (z.B. GRIS)
  - Grid Information Protocol (GRIP)
    => Benutzer
  - Grid Registration Protocol (GRRP)
    => GIIS
- Aggregate Directory Services (z.B. GIIS)
  - Grid Information Protocol (GRIP)
    => Benutzer
  - Grid Registration Protocol (GRRP)
    => GIIS
- => Hierarchien

MDS

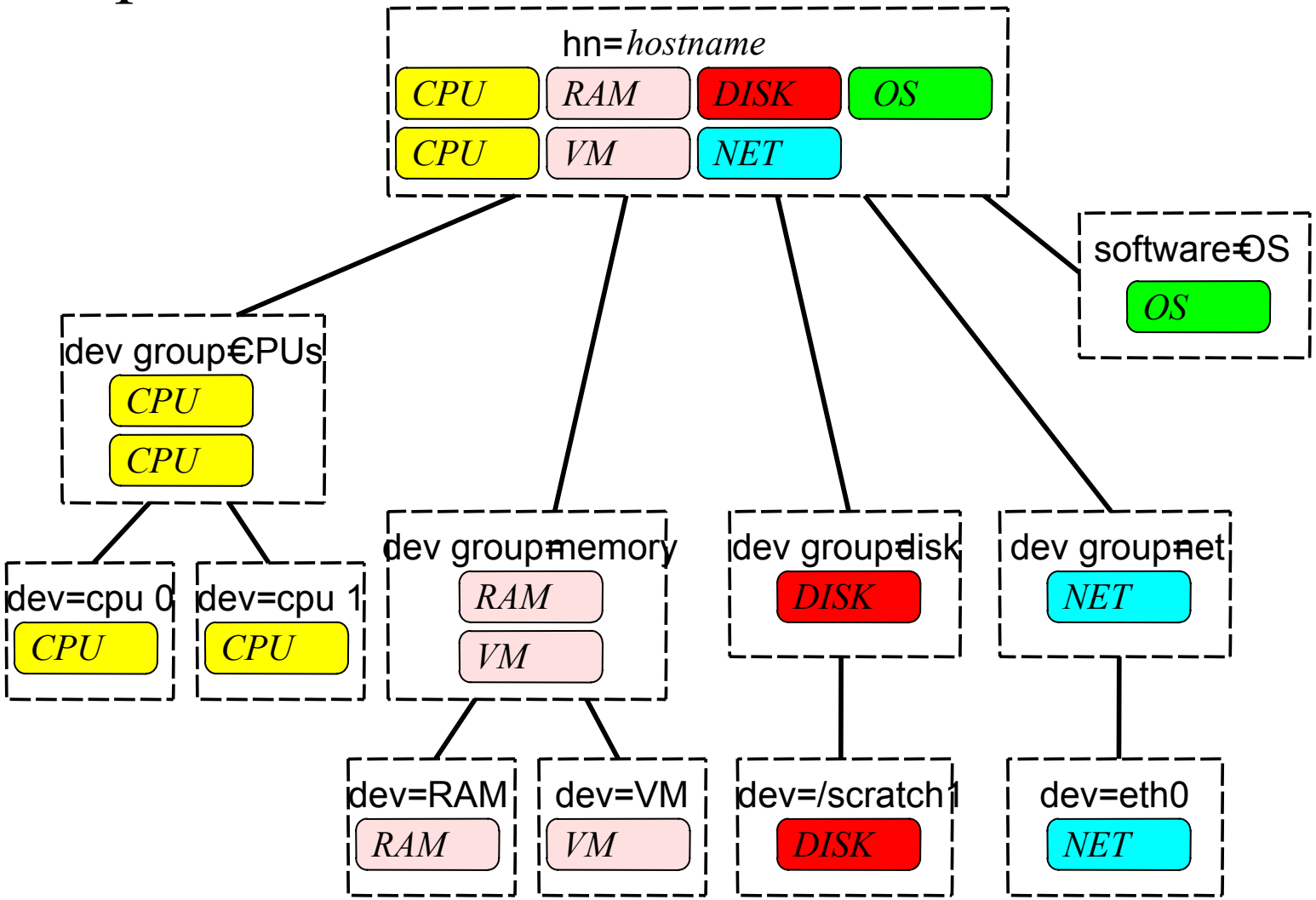
- Inhalt
  - Ressource Informationen (Hardware, Jobmanager)
    - Quelle: GRAM Reporter
  - Netzwerk Informationen
    - Quelle: z.B. NWS
  - Informationen zur Software-Ausstattung
  - Aktualität: Zeitstempel

MDS


Client 1
Client 2
Client 3
Resource A
GRIS
Resource B
GRIS
Clients 1 and 2 request info directly from resources.
Client 3 uses GIIS for searching collective information.
GIIS requests information from GRIS services as needed.
GIIS
Cache contains info from A and B

MDS

- Beispiel: Host



GridFTP

- Dient zur Übertragung großer Datenmengen zwischen Grid Knoten
- FTP Server
- Erweitert um
  - GSI Authentifizierung
  - Striped Data Transfer
  - Third Party Transfer
  - Partial Data Transfer
- Optimiert fuer große Datenmengen
- Interaktiv oder Programmgesteuert

GASS

Global Access to Secondary Storage

- Transparenter Zugriff auf entfernte Dateien
- Bietet einen Caching Mechanismus
- Automatisches Versenden von Binaries
- Protokolle:
  - HTTP
  - FTP
  - ...
- Aber: limitiert auf kleine Datenmengen
  Sonst: GridFTP

Globus IO

- Wrapper für BSD Sockets
- Vergleichbare Programmierung
- Optionaler GSI-Support

Benutzerperspektive

Verwendung des Grids (Vorbereitung):

1. Umgebung setzen

   $GLOBUS_LOCATION/etc/globus_user_env.(c)sh

2. Benutzerzertifikat beantragen

   grid-cert-request

   Output => ca@vcpc.univie.ac.at (Istzustand)

   Erhaltenes Zertifikat in ~/.globus/usercert.pem speichern

   Eintragung im grid-mapfile

Benutzerperspektive

- Verwendung:

1. Umgebung setzen

   $GLOBUS_LOCATION/etc/globus_user_env.(c)sh

2. Benutzerproxy starten

   grid-proxy-init

3. Testaufruf

   globus-job-run <host> /bin/date

4. Benutzerproxy entfernen

   grid-proxy-destroy

Benutzerperspektive

- Weitere Kommandos (bei aktivem proxy)
  - grid-proxy-info (z.B. verbleibende Gültigkeit)
  - grid-info-search –x –h <host> (MDS)
  - globus-url-copy gsiftp://localhost/tmp/file1 file:///tmp/file2 (GridFTP)

Benutzerperspektive

- Hintergrundjobs
  - globus-job-submit <host> /bin/hostname
    - => liefert URL
  - globus-job-status <url>
  - globus-job-get-output <url>
  - globus-job-clean <url>
  - globus-job-cancel <url>

RSL

- Dient zur genaueren Spezifikation der Aufgabe
- Resource Specification Language
  - globusrun -o –r <host> `&(executable=/bin/ls)(arguments=„-l -a“)`
- Üblich: RSL-Datei
  - globusrun –w –f <Datei>

RSL

RSL-Beispiel: MPI-Job (DUROC-Multirequest)

```
+
(&(resourceManagerContact=„iris“)
      (count=2)
      (environment=(GLOBUS_DUROC_SUBJOB_INDEX 0)
                        (LD_LIBRARY_PATH /opt/globus/lib))
      (directory=„/home/gup/ph/globus_test“)
      (executable=„/home/gup/ph/globus_test/ring“)
)
(&(resourceManagerContact=„pan“)
      (count=2)
      (environment=(GLOBUS_DUROC_SUBJOB_INDEX 1)
                        (LD_LIBRARY_PATH /opt/globus/lib))
      (directory=„/home/gup/ph/globus_test“)
      (executable=„/home/gup/ph/globus_test/ring“)
)
```

Programmierung

- Modulare Struktur
  - Infrastruktur
  - Kommandos
  - API
- Zuordnung zu den Globus Diensten
  - Common Programming Infrastrcuture
  - GRAM
  - MDS
  - Security
  - Globus IO
  - GridFTP
  - GASS

Common Programming Infrastructure

- Grundlegendes Modul
- Wird vom Rest der Module verwendet
- Aktivierung/Deaktivierung von Modulen
- Thread-Bibliothek (Posix Teilmenge)
- Thread-sichere und portable libc-Wrapper
- Callback Routinen
- Datenobjekt- und Fehlermanagement
- Elementare Datenstrukturen (FIFOS, Listen, ...)

Common Programming Infrastructure

- globus_module_activate(GLOBUS_IO_MODULE)
- globus_module_deactivate(GLOBUS_IO_MODULE)
- Automatisches Laden von Abhängigkeiten
- Mehrfaches Laden möglich

Common Programming Infrastructure

Threads

- globus_thread_*(), globus_mutex_*(), globus_cond_*()
- Einfache POSIX threads (pthreads) Teilmenge
- Selbe Parameter und Semantik wie pthreads
- Änderung: “pthread” zu “globus” oder “globus_thread” im Functionsnamen
- Portabilität zu nicht-pthread basierten Systemen
- Arbeitet mit Programmen zusammen, die pthreads direkt nutzen

Common Programming Infrastructure

Erzeugung und Management von Threads:

- globus_thread_create()
- globus_thread_exit()
- globus_thread_self()
- globus_thread_equal()
- globus_thread_once()
- globus_thread_yield()
- globus_threadattr_init(), …_destroy()
- globus_threadattr_setstacksize(), …

Common Programming Infrastructure

- Mutual Exclusion
  - globus_mutex_init()
  - globus_mutex_destroy()
  - globus_mutex_lock()
  - globus_mutex_trylock()
  - globus_mutex_unlock()
- Weitere Möglichkeiten:
  - Conditions
  - Callbacks

GRAM

- Ressource Management APIs: RSL, GRAM, DUROC
  - globus_rsl
  - globus_gram_client
  - globus_gram_myjob
  - globus_duroc_control
  - globus_duroc_runtime

GRAM

Globus RSL

- Modul für die Manipulation von RSL Ausdrücken
  - Parsen von RSL Zeichenketten in eine Datenstruktur
  - Manipulation der Datenstruktur
  - Ausgabe der Datenstruktur in eine Zeichenkette
- Kann zur Programmierung von Brokers dienen, die RSL-Spezifikationen verfeinern

GRAM

globus_gram_client

- Dient zur Programmgesteuerten Ausführung von remote jobs
- globus_gram_client_job_request()
- globus_gram_client_job_status()
- globus_gram_client_job_cancel()

MDS

- MDS Client Programmierung:
  LDAP Protokoll
- z.B. OpenLDAP Client Bibliothek
  - ldap_open()
  - ldap_close()
  - ldap_search_s()
  - ldap_first_entry()
  - ldap_next_entry()
  - ldap_first_attribute()
  - ldap_next_attribute()
  - ldap_get_values()

Security

globus_gss_assist

- Das globus_gss_assist Modul ist ain Globus Toolkit spezifischer Wrapper um die GSS-API
  - Versteckt einige Details der GSS-API
  - Hält sich an Globus Toolkit Konventionen
  - Trennung von der Kommunikationsmethode

Globus IO

Motivation

- Zahlreiche Anwendungen benutzen Kombinationen von TCP, UDP, IP Multicast, und Datei E/A.
  - "Neuerfindung des Rades"
  - Selten Sicherheitsaspekte berücksichtigt
- Vorteile von globus_io
  - Einfach: Sicherheit, Socket Optionen (z.B. non-blocking)
  - Very similar to existing BSD socket calls

Globus IO

Ansatz:

- Socket- und Dateiabstraktionen
- Synchrone und Asynchrone Versionen
- Sicherheit und Socket Optionen, durch Attribute
- I/O Handles repräsentieren die Dateien oder Verbindungen

Globus IO

Gemeinsame Funktionen für alle Kommunikationsarten

- globus_io_[register]_select()
- globus_io_[register]_cancel()
- globus_io_[register]_close()
- globus_io_[register]_read()
- globus_io_[register]_write()
- globus_io_[register]_writev()
- globus_io_try_{read,write,writev}()
- globus_io_get_handle_type()
- globus_io_handle_{set,get}_user_pointer()

Globus IO

- TCP-Verbindung
  - globus_io_tcp_create_listener()
  - globus_io_tcp_[register]_listen()
  - globus_io_tcp_[register]_accept()
  - globus_io_tcp_[register]_connect()
- Attribute
  - globus_io_tcp_set_attr()
  - globus_io_tcp_get_attr()

Grid FTP

APIs

- globus_ftp_control
  - Low-level GridFTP Kontroll- und Datenkanal Operationen
- globus_ftp_client
  - Clientoperationen
- globus_gass_copy
  - Datentransfer über GridFTP, HTTP, lokale Datei und Speicher

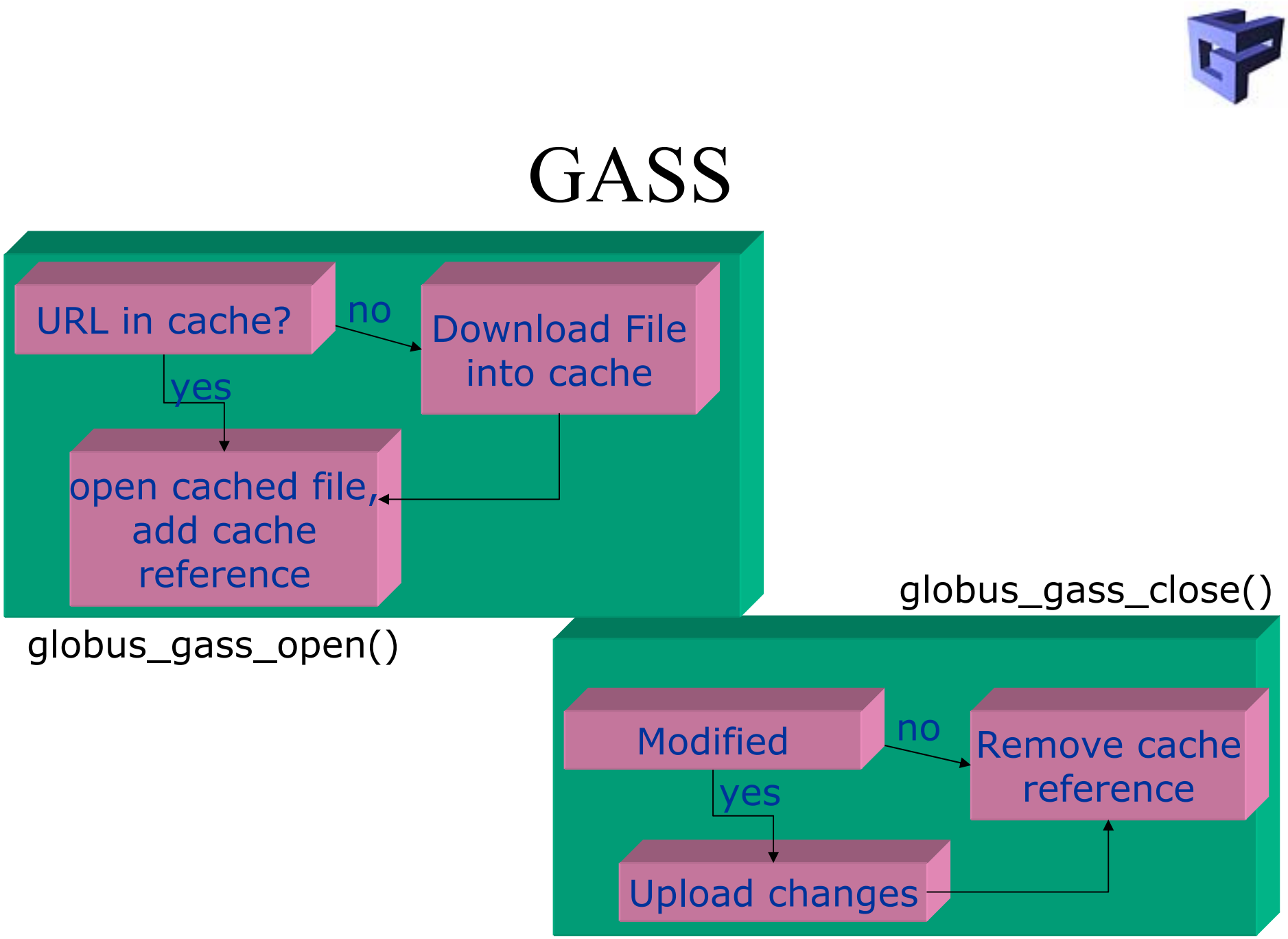
GridFTP

- Verwaltung über Handles mit Attributen
- globus_ftp_client_*
  - globus_ftp_client_get
  - globus_ftp_client_put
- Spezielle GridFTP features werden über Handleattribute gesteuert
  - globus_ftp_client_operationattr_set_<attribute>(&attr, &<attribute_struct>)
  - globus_ftp_client_operationattr_get_<attribute>(&attr, &<attribute_struct>)

GASS

Dateizugriffs-API

- Dateitransfer über GASS-Server, wird bei Bedarf gestartet
- Minimale Änderungen im Programm
- globus_gass_open(), globus_gass_close()
  - Wie open(), close() aber URLs statt Dateinamen
  - URL (Datei) wird lokal gecacht für mehrfaches Öffnen
  - Deskriptoren zu lokalen Dateien oder Sokets für remote Verbindungen
  - globus_gass_fopen(), globus_gass_fclose()


GASS
URL in cache?
no
Download File into cache
yes
open cached file, add cache reference
globus_gass_open()
globus_gass_close()
Modified
no
Remove cache reference
yes
Upload changes

GASS

globus_gass_cache

- Modul für die Manipulation des GASS Cache
- globus_gass_cache_open(), …_close()
- globus_gass_cache_add(), …_add_done()
- globus_gass_cache_delete_start(), …_delete()
- globus_gass_cache_cleanup_tag()
- globus_gass_cache_cleanup_file()
- globus_gass_cache_list()
- Kein Füllen des Inhalts. Management von Namen und Lebenszeit

GASS

globus_gass_transfer

- API zum Transport von remote remote Dateien über verschiedene Protokolle
  - http
  - https
- Put und Get Operationen auf einen URL
- Effizienter Transfer zwischen Dateien oder direkt von/zum Speicher

GASS

globus_gass_copy

- Einfache API zum Kopieren von Dateien zwischen Quelle und Ziel
- URL für Quellen- und Zeilangabe
- http(s), (gsi)ftp, file
- Von ftp zu ftp 3rd party Transfer