
InGriDE: Um Ambiente Integrado de Desenvolvimento para Computação em Grade

Eduardo Guerra
eguerra@ime.usp.br

Orientador: Prof. Dr. Alfredo Goldman

Proposta de dissertação apresentada ao IME-USP para qualificação de Mestre em Ciência da Computação

Fevereiro, 2006

Visão Geral

- Cenário ideal
- Realidade atual
- Nossa proposta
 - IDE para Computação em Grade
 - Portalidade
 - Aplicações escritas
 - Ferramentas do ambiente

Agenda

- Computação em Grade
- Ferramentas para Computação em Grade
- Definição do Problema
- InGriDE
- Plano de Trabalho

Computação em Grade

- Surgimento
 - Crescimento da Internet
 - Avanços tecnológicos e redução de custos
 - Uso de computadores geograficamente distribuídos
- Desafios
 - Surgiram em diferentes níveis (eg. modelos conceituais e de implementação, projeto e desenvolvimento de aplicações, segurança)

Computação em Grade (cont.)

- Desenvolvimento dos **Sistemas de Computação em Grade** ou **middlewares**
- Movimento tanto no meio acadêmico quanto na indústria
- Definição: “*Uma Grade Computacional é uma infra-estrutura de software capaz de interligar e gerenciar diversos recursos distribuídos geograficamente, com o objetivo de oferecer ao usuário da Grade acesso transparente a estes recursos, independente de sua localização*” Goldchleger, 2004

Computação em Grade (cont.)

- Diversidade de projetos --> falta de padronização
- Iniciativa do Global Grid Forum (GGF)
 - Padrão Open Grid Service Architecture (OGSA)/
Open Grid Service Infrastructure (OGSI)
 - Padrão Web Services Resource Framework
(WSRF)
 - Grupos de pesquisa em diferentes áreas

Computação em Grade (cont.)

■ *Middleware*s:

□ Tradicionais

- **Globus**: GT3 implementa OGSA e GT4 implementa WSRF
- **Legion**: modelo OO elegante e flexível
- **MyGrid/OurGrid**: foco em aplicações Bag-of-Tasks(BOTs)
- **ProActive**: biblioteca como uma camada de alto nível

□ Oportunistas

- **Condor**: mais antigo, suporte a aplicações paralelas, *checkpointing*
- **BOINC**: sucessor de Setti@Home

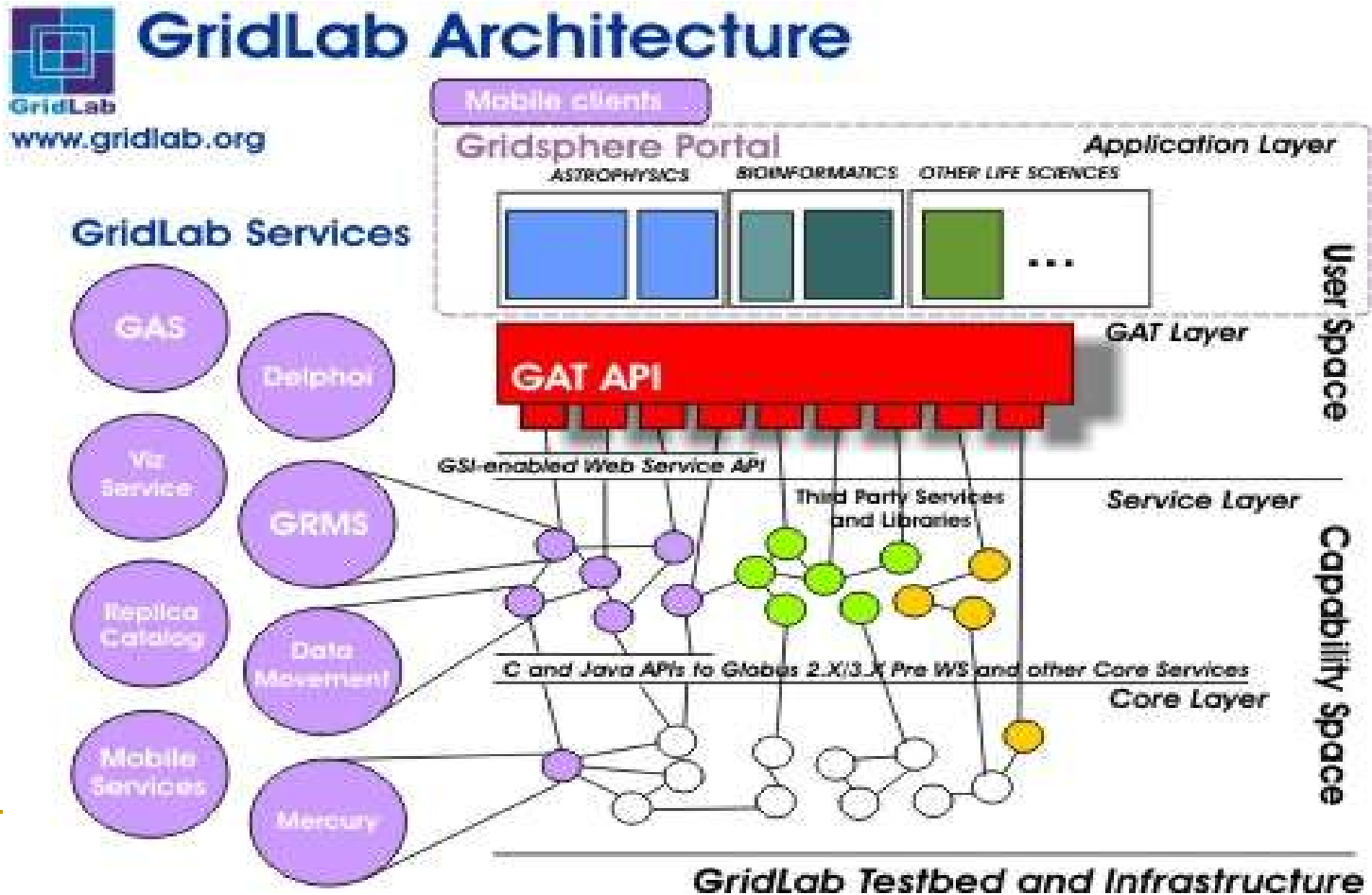
Computação em Grade (cont.)

- GridLab

- *Middleware* com diferentes serviços de Grade
- Espécie de integrador de Grades: possibilita a utilização de serviços de terceiros
- Comunidade: instituições da Europa e dos Estados Unidos incluindo Laboratório de Argonne em Illinois, Universidade de Cardiff, HP e Sun
- Principal Objetivo: fornecer uma camada que simplifique o acesso das aplicações ao *middleware*

Computação em Grade (cont.)

- Grid Application Toolkit (GAT)



Ferramentas para Computação em Grade

- O que são?
- Atuação em diferentes pontos da Grade
 - Lado servidor x lado cliente x ambos
- *Nosso foco: **Grid Computing Environments** (GCEs) ou **Grid Application-level Tools***
 - Abstração da Grade no lado cliente
- Os GCEs normalmente oferecem:
 - Ambientes de programação
 - Ferramentas para interação com a Grade

Ferramentas para Computação em Grade

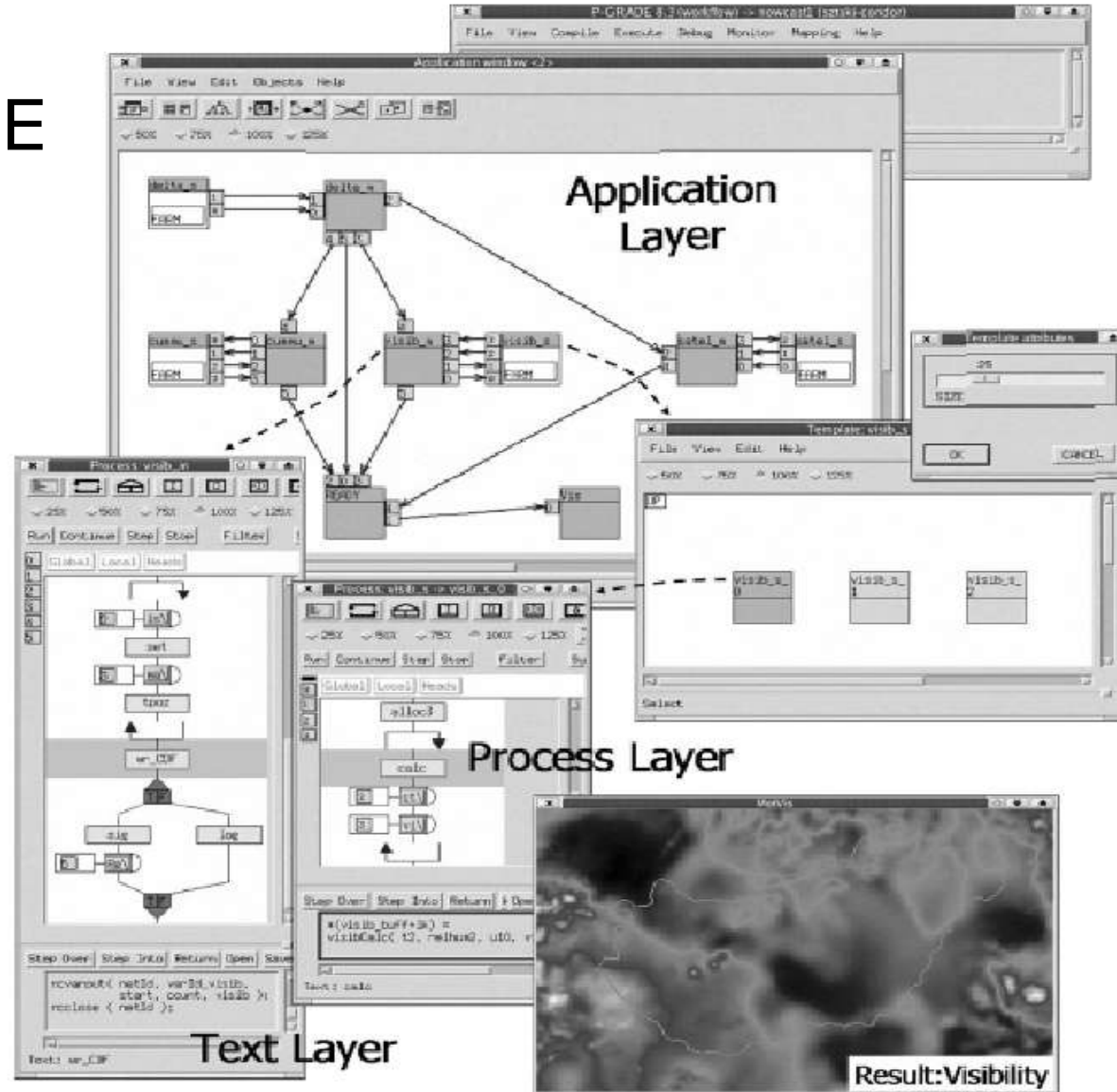
- GCEs com diferentes níveis de abstração
 - Arquitetura em camadas
- Classificação em grupos:
 - *Workflows*
 - *Problem Solving Environments (PSEs)*
 - *Parameter Sweep Applications (PSAs)*
 - Portais
 - Outras

Ferramentas - *Workflows*

- Construção de aplicações usando componentes gráficos
- Importância: aplicações não são mais monolíticas (indústria e meio acadêmico)
- Adequação à arquitetura da Grade
- Linguagens
 - Específicas para Grade, mas baseadas em linguagens de composição de Web Services
- Problema da Padronização: Grupo do GGF

Ferramentas - Workflows

- P-GRADE

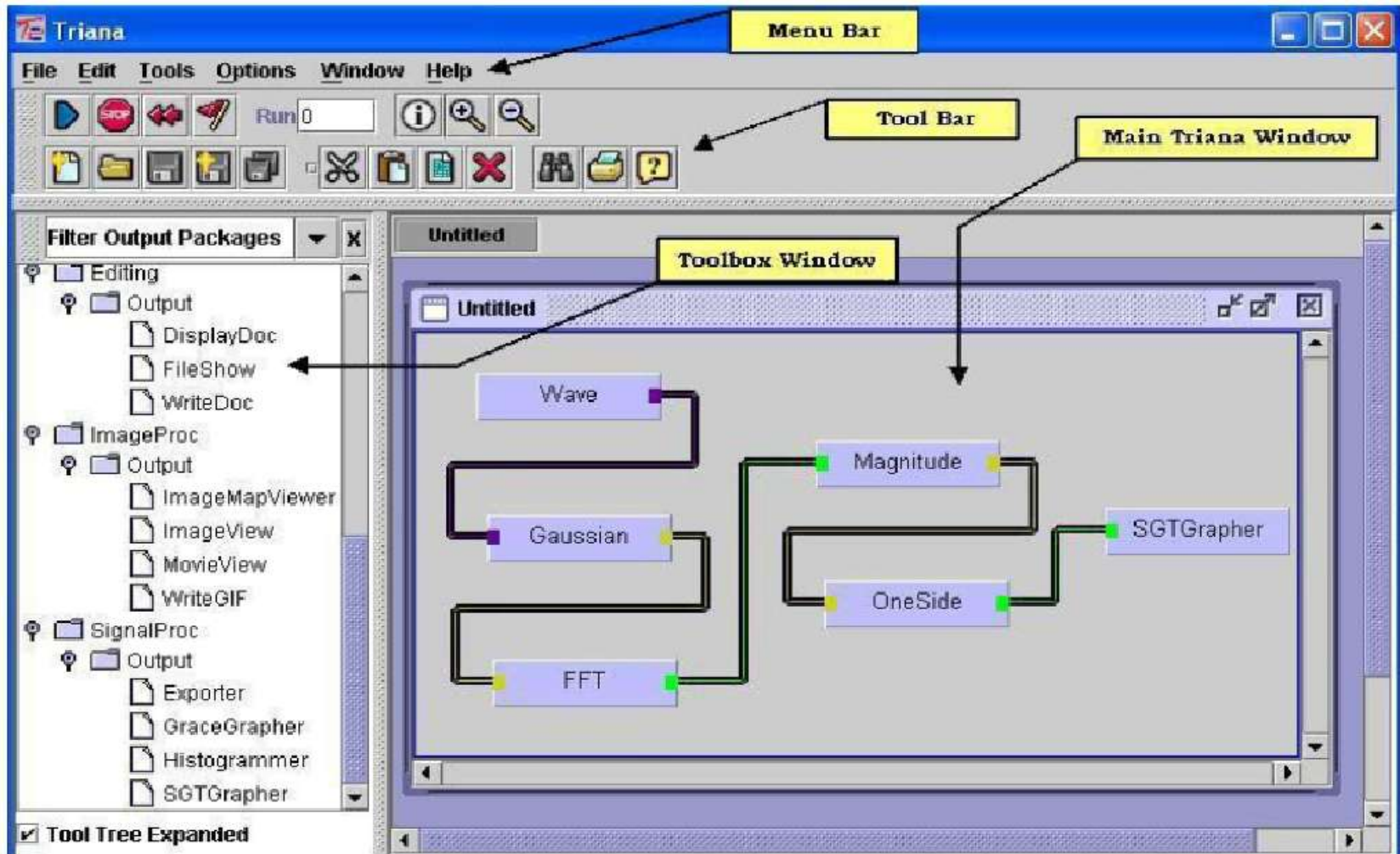


Ferramentas - *Problem Solving Environments*

- Construção de aplicações para um domínio específico
- Fornece componentes reutilizáveis
- Normamente dependentes de uma Grade Computacional
- Eficiente na disseminação da Computação em Grade junto à comunidade de usuários

Ferramentas - Problem Solving Environments

■ Triana

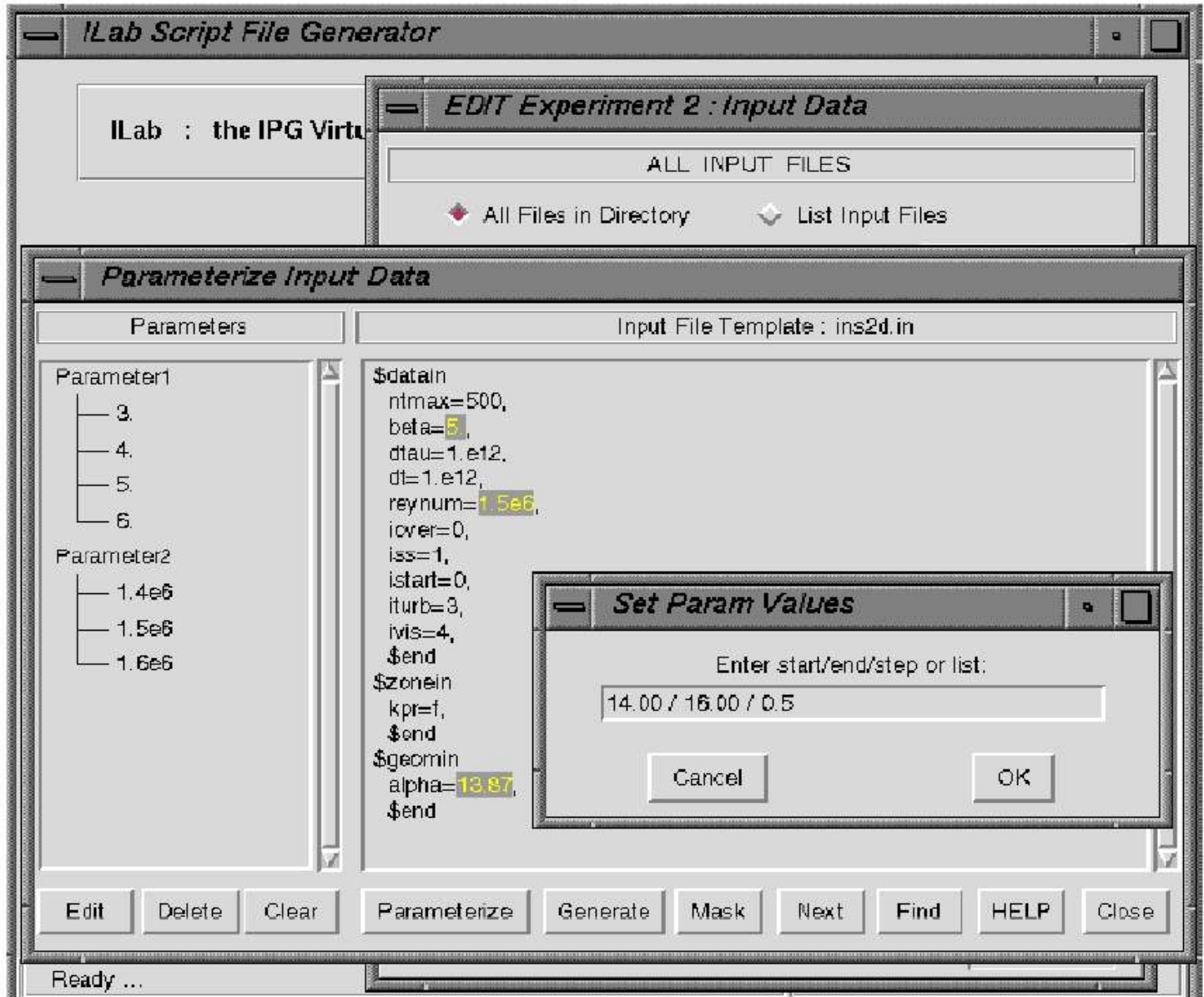


Ferramentas – *Parameter Sweep Applications*

- Tipo específico de aplicação
 - Tarefas independentes
 - Nós recebem conjunto de dados diferentes
 - Grande comunidade
- Modelo simplificado favorece o uso da Grade
 - Mais tolerante à latência de rede
 - Mais suscetíveis a mecanismos de tolerância a falhas
- Situação de execução: ***Estudo Paramétrico***

Ferramentas – *Parameter Sweep Applications*

- ILab



Ferramentas - Portais

- Uso do termo **Portal**
- **Portal de Grade (Grid Portal)**
 - Informações sobre o estado de recursos e serviços
 - Interação com os serviços da Grade de forma simplificada em um ambiente (Web) conhecido
- Três gerações destacadas por Li e Baker:
 - Primeira: foco nas funcionalidades e no Globus
 - Segunda: uso de **Portlets** (personalização)
 - Terceira: **Portlets** com anotações semânticas

Ferramentas - Portais

■ GridSphere

GridSphere Portal - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

CCT Portal

Welcome | Grid

Credentials Resources Jobs Files

Logout
Welcome, Oliver Wehrens

Resource Browser Portlet

<< Resource List Refresh View

Peyote

Hostname	peyote.aei.mpg.de	CPU Count	4
Description	Front-end to the AEI Peyote Cluster	CPU Cache	512 X 12kb
Platform	i686	CPU Speed	2799Mhz
CPU Vendor	GenuineIntel	CPU Load 1 min	285
CPU Version	15.2.7	CPU Load 5 min	277
CPU Model	Intel(R) Xeon(TM) CPU 2	CPU Load 15 min	276
CPU Features	fpu vme de pse tsc mce sse sse2 mmx fxsr	Memory Size	4038MB
OS Name	Linux	Memory Free	3608MB
OS Release	2.4.25	Virtual Memory Size	1027MB
		Virtual Memory Free	296MB

Services, Accounts, JobQueue information are available.

Resource Profile **Services** View

	Protocol	Port	Service Path
GRIS		2135	
Globus Gatekeeper		2119	
Grid Ftp		2811	

Ferramentas – Outras

- GT4IDE
 - Usa o arcabouço Eclipse
 - Específico para o Globus
 - Desenvolvimento de **serviços** para a Grade
- Gride
 - Usa o arcabouço Netbeans
 - Específico para o Globus
 - Ferramentas para o ciclo completo de desenvolvimento de **aplicações**

Definição do Problema

- Situação atual:
 - Usuários utilizam várias ferramentas não integradas
 - Essas ferramentas são específicas para um *middleware*
 - As aplicações escritas não são portáteis para diferentes middlewares
- Carência de uma ferramenta integrada que atenda às necessidades dos usuários
- Disseminação da Computação em Grade

InGriDE

- ***Integrated Grid Development Environment***
 - IDE para Computação em Grade
 - Atender o ciclo completo de desenvolvimento
 - *Middleware* InteGrade usado como estudo de caso
- Tipos de usuários da Grade
 - Desenvolvedores de aplicações (InGriDE e Portal)
 - Usuários finais (Portal)
- Requisitos não-funcionais
 - Facilidade de uso
 - Portabilidade em relação ao *middleware*

InGriDE

- Objetivos e Desafios
 - Levantamento das funcionalidades necessárias para o desenvolvedor no ciclo de desenvolvimento
 - Integração dessas funcionalidades em uma única ferramenta fácil de usar
 - Portabilidade em relação ao *middleware*
 - Ferramenta
 - Aplicações escritas
 - Adaptação das necessidades do *middleware* InteGrade à solução de portabilidade

InGriDE

■ InteGrade

- ❑ *Middleware* orientado a objetos para Computação em Grade Oportunista
- ❑ Comunidade: USP, UFMS, PUC-Rio UFMA e UFG
- ❑ Visa manter a qualidade de serviço das estações
- ❑ Usa o padrão CORBA na interligação dos objetos
- ❑ Suporte a aplicações paralelas baseadas no modelo BSP
- ❑ Mecanismo de *checkpointing*
- ❑ Coleta e análise de padrões de uso

InGriDE

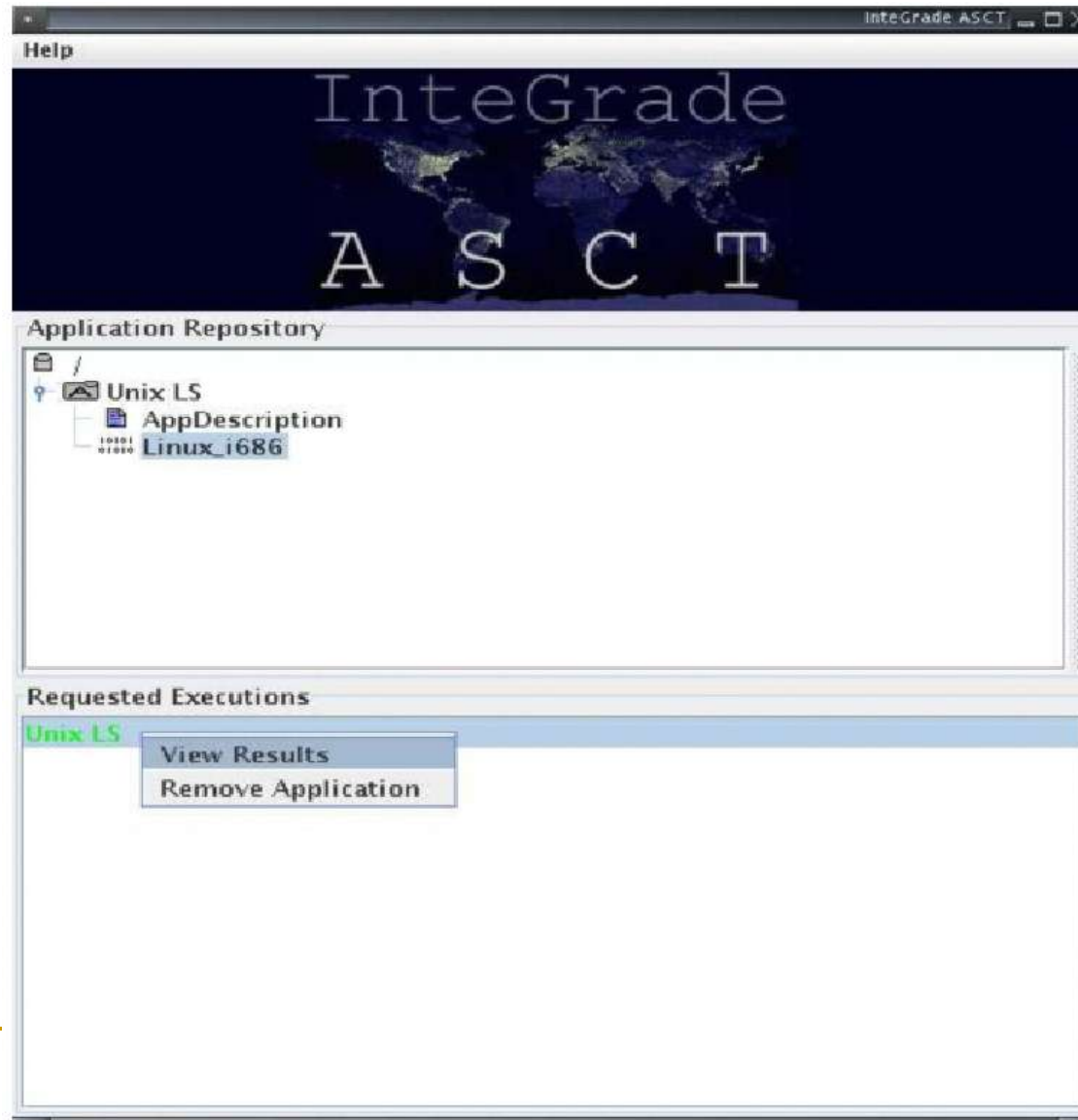
- **InteGrade**

- Ferramentas existentes

- *Application Submission and Control Tool (ASCT)*
 - ClusterView

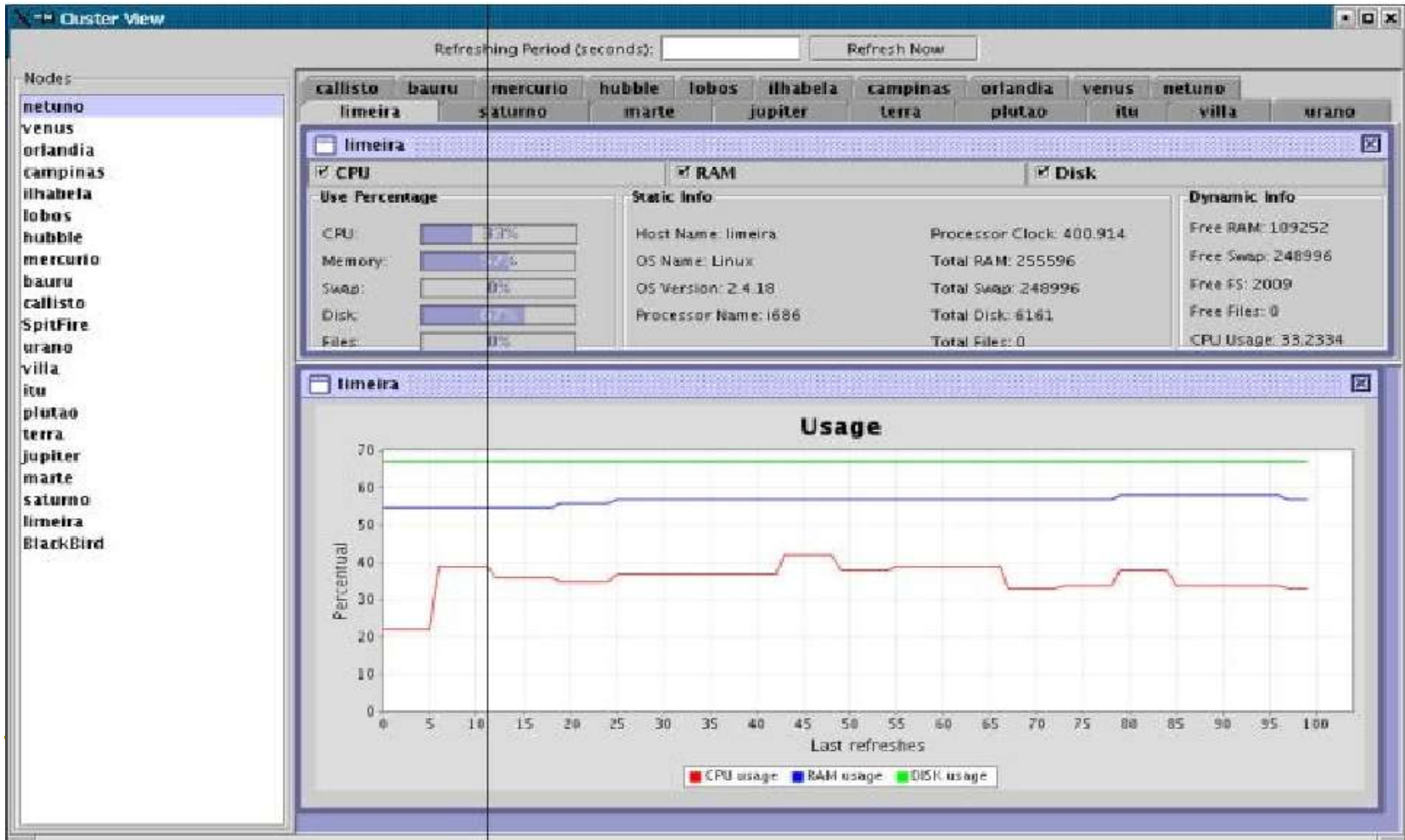
InGriDE

- ASCT



InGriDE

ClusterView



InGriDE

- Solução
 - Levantamento de requisitos
 - Ferramentas encontradas na literatura
 - Necessidades dos usuários do InteGrade
 - Integração das ferramentas
 - Arcabouço Eclipse
 - Portabilidade
 - Inspiração no *Grid Application Toolkit* (GAT) do projeto GridLab

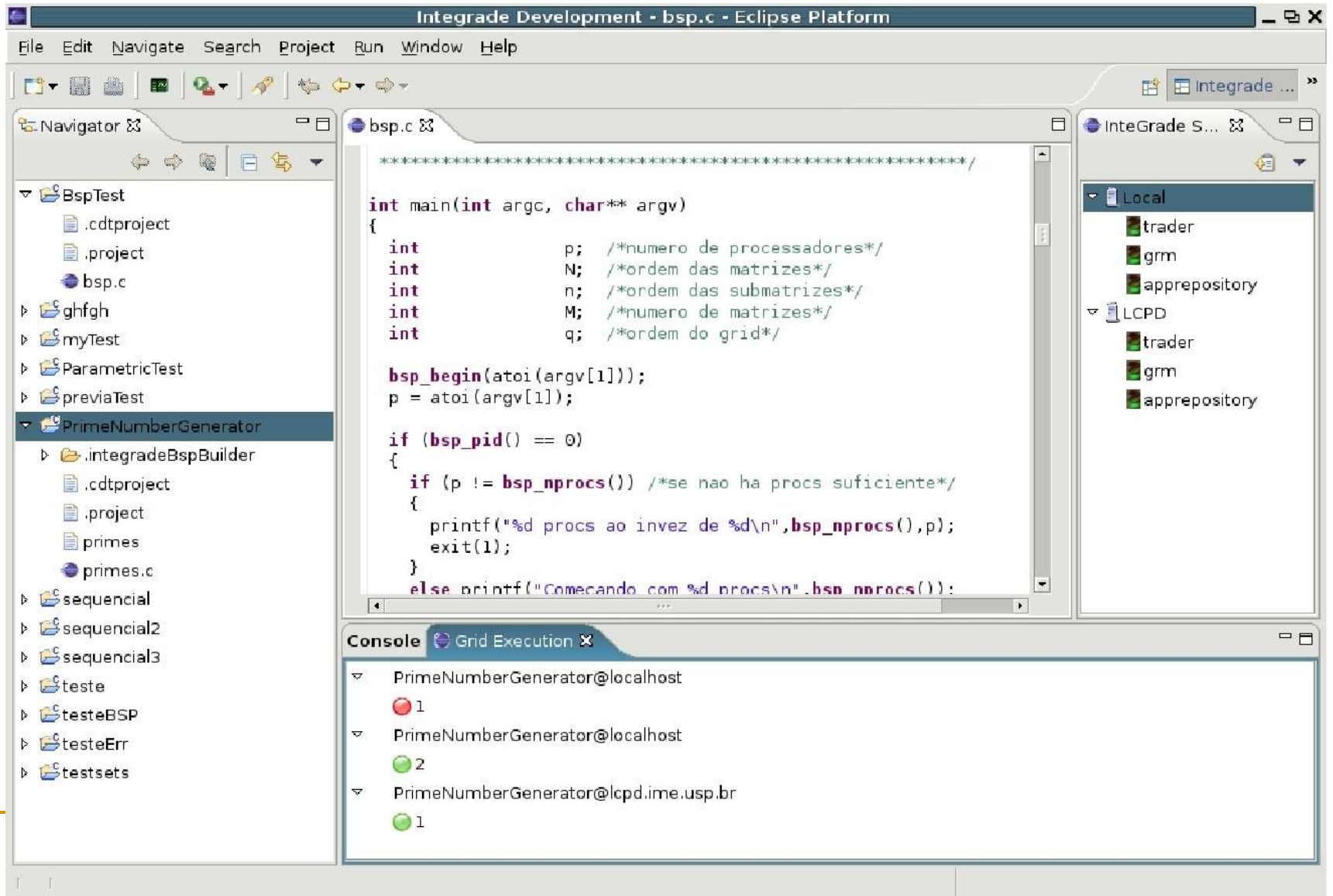
InGriDE

- Funcionalidades
 - ✓ Submissão de tarefas
 - ✓ Monitoramento da execução das aplicações
 - ✓ Interação com o ambiente de execução
 - Busca de recursos
 - Editor de código com verificação de consistência
 - Obtenção de estatísticas de execução
 - Facilidades para simulações

InGriDE

- Funcionalidades
 - Visualização das informações de padrão de uso
 - Instrumentação de código para *checkpointing*
 - Interação com a ambiente de segurança

InGriDE



Integrade Development - bsp.c - Eclipse Platform

File Edit Navigate Search Project Run Window Help

Navigator

- BspTest
 - .cdtproject
 - .project
 - bsp.c
- ghfgh
- myTest
- ParametricTest
- previaTest
- PrimeNumberGenerator**
 - integradeBspBuilder
 - .cdtproject
 - .project
 - primes
 - primes.c
 - sequencial
 - sequencial2
 - sequencial3
 - teste
 - testeBSP
 - testeErr
 - testsets

bsp.c

```
*****/
int main(int argc, char** argv)
{
    int p; /*numero de processadores*/
    int N; /*ordem das matrizes*/
    int n; /*ordem das submatrizes*/
    int M; /*numero de matrizes*/
    int q; /*ordem do grid*/

    bsp_begin(atoi(argv[1]));
    p = atoi(argv[1]);

    if (bsp_pid() == 0)
    {
        if (p != bsp_nprocs()) /*se nao ha procs suficiente*/
        {
            printf("%d procs ao inves de %d\n", bsp_nprocs(), p);
            exit(1);
        }
        else printf("Comecando com %d procs\n", bsp_nprocs());
    }
}
```

InteGrade S...

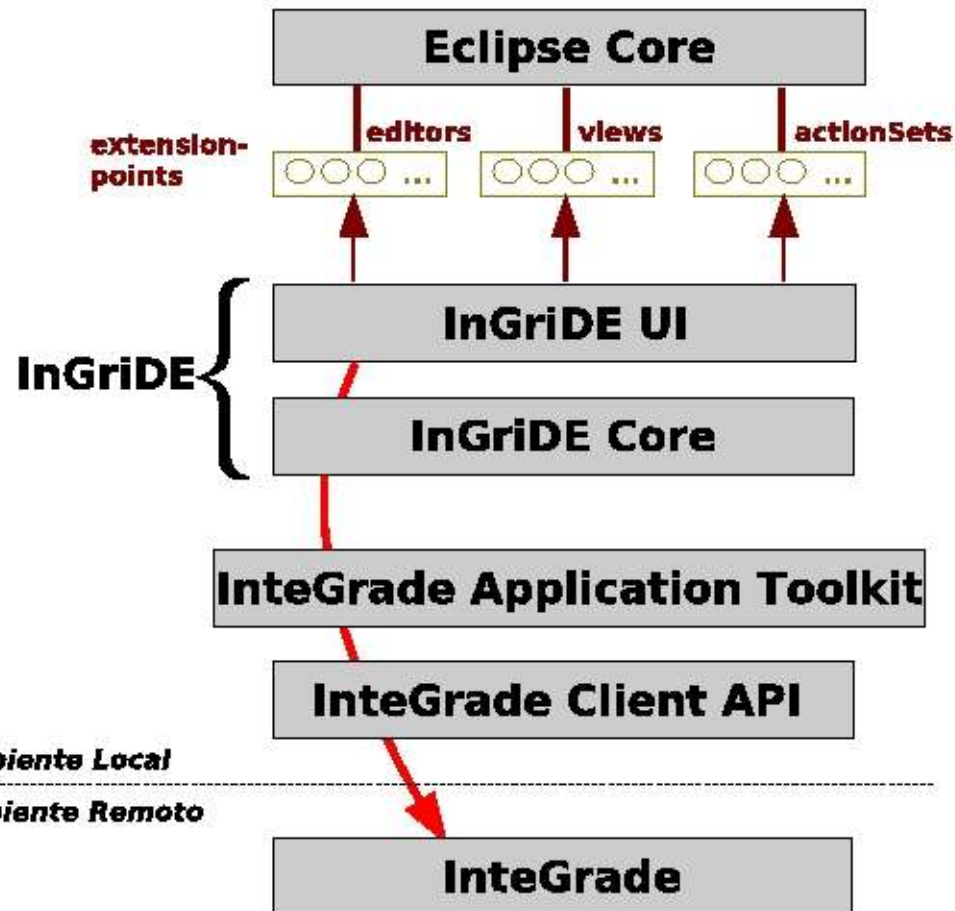
- Local
 - trader
 - grm
 - apprepository
- LCPD
 - trader
 - grm
 - apprepository

Console Grid Execution

- PrimeNumberGenerator@localhost
 - 1
- PrimeNumberGenerator@localhost
 - 2
- PrimeNumberGenerator@lcpd.ime.usp.br
 - 1

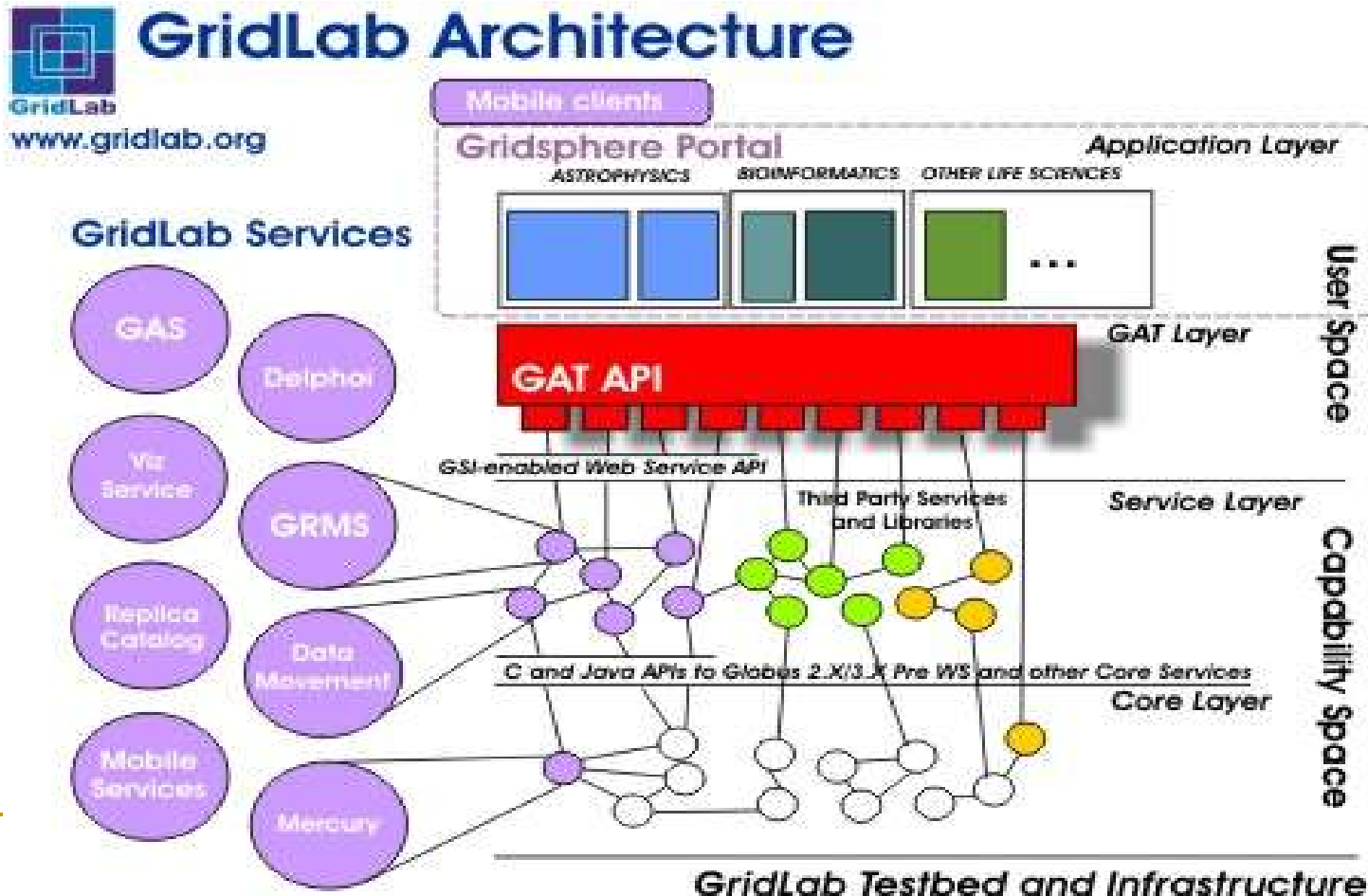
InGrIDE

- Arquitetura - Visão Geral



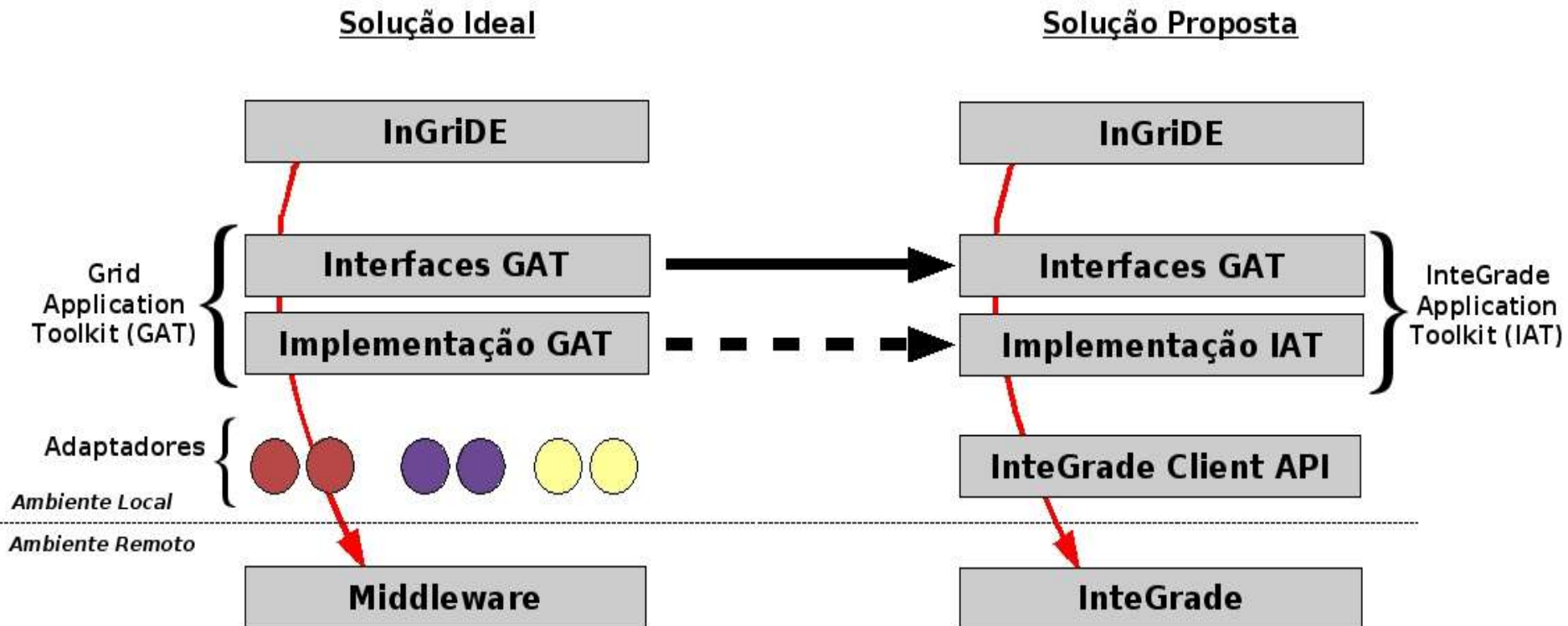
InGriDE

- Arquitetura – Portabilidade baseada em GAT



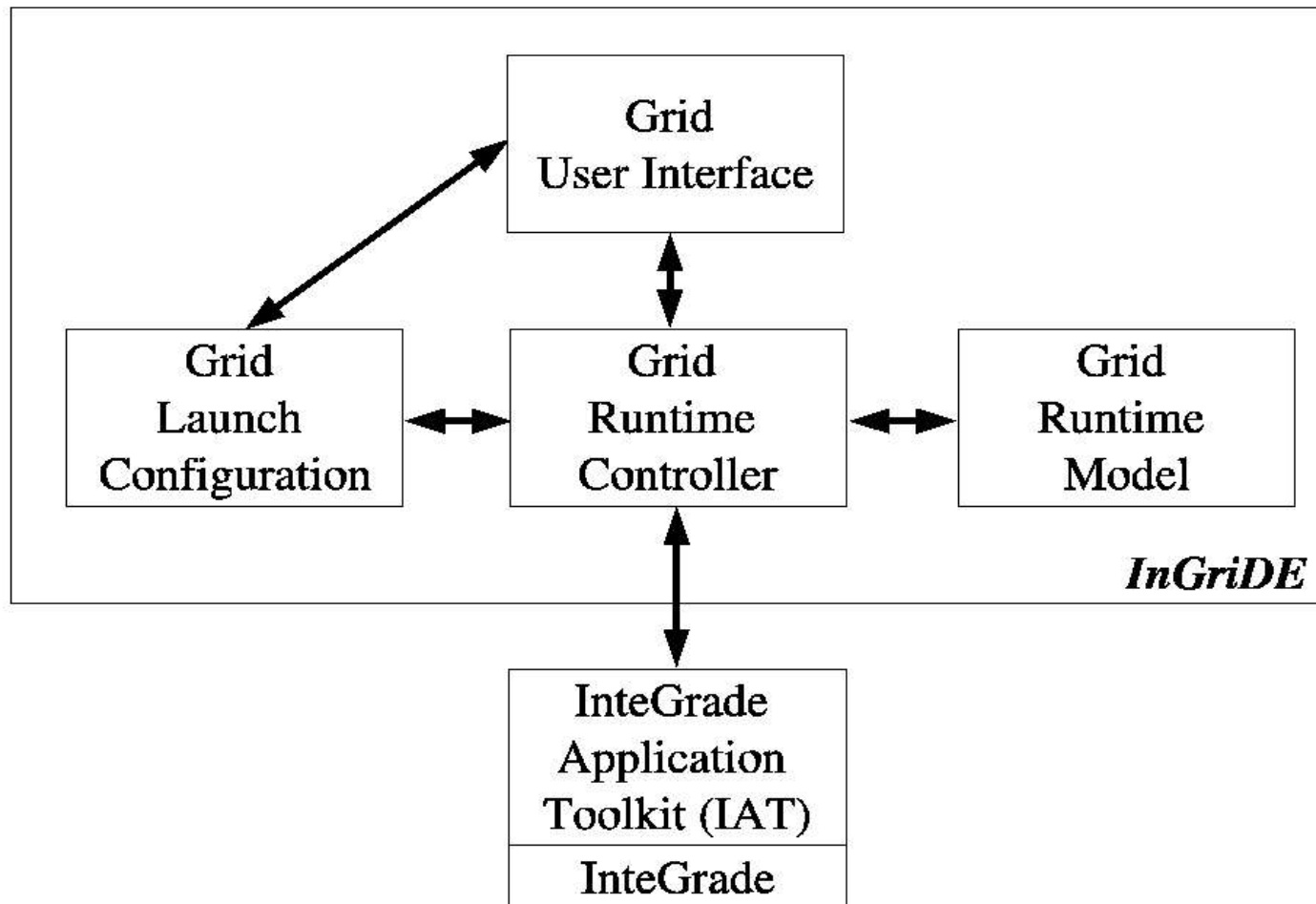
InGriDE

- Arquitetura – Portabilidade baseada em GAT



InGriDE

- Arquitetura - Projeto interno



Resumo

- Computação em Grade
- Ferramentas para Computação em Grade
- Problema
- InGriDE

Plano de Trabalho

- Primeira etapa (3 meses)
 - Adequar a implementação atual à arquitetura proposta
 - Disponibilizar uma versão testada para o grupo InteGrade (*feedback*)
- Segunda etapa (3 meses)
 - Levantamento e inclusão de novas funcionalidades
 - Criação de um Portal usando GridSphere
 - Redação da dissertação

Plano de Trabalho (Cronograma)

Atividade	Fev	Mar	Abr	Mai	Jun	Jul
Construção da InteGrade Application Toolkit	x	x				
Adequação da implementação atual à arquitetura proposta		x				
Melhorias nas funcionalidades atuais		x	x			
Disponibilização de uma versão para o projeto InteGrade			x			
Criação de um projeto na Incubadora FAPESP			x			
Desenvolvimento de novas funcionalidades				x	x	
Construção do Portal InteGrade baseado no GridSphere				x	x	
Redação de artigo para conferência internacional						x
Redação da dissertação				x	x	x

InGriDE: Um Ambiente Integrado de Desenvolvimento para Computação em Grade

Eduardo Guerra
eguerra@ime.usp.br

Orientador: Prof. Dr. Alfredo Goldman

Proposta de dissertação apresentada ao IME-USP para qualificação de Mestre em Ciência da Computação

Fevereiro, 2006