



# SymGrid: a Framework for Symbolic Computations on the Grid

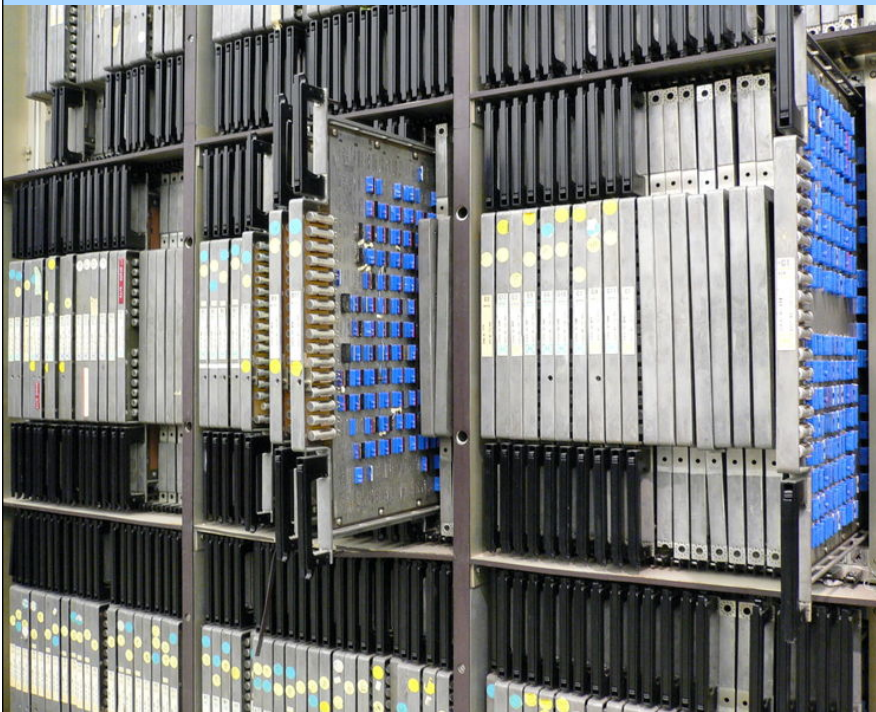
**Abdallah Al Zain**  
**Heriot-Watt University, Scotland**

## Overview

1. Introduction
  - Background and Motivation
2. The SymGrid System
  - SymGrid Objectives
  - SymGrid-Services
  - SymGrid-Par
3. *Recent* Single-Cluster and Multicore Results
4. Linking SymGrid with Web Services
5. Conclusions and Further Work

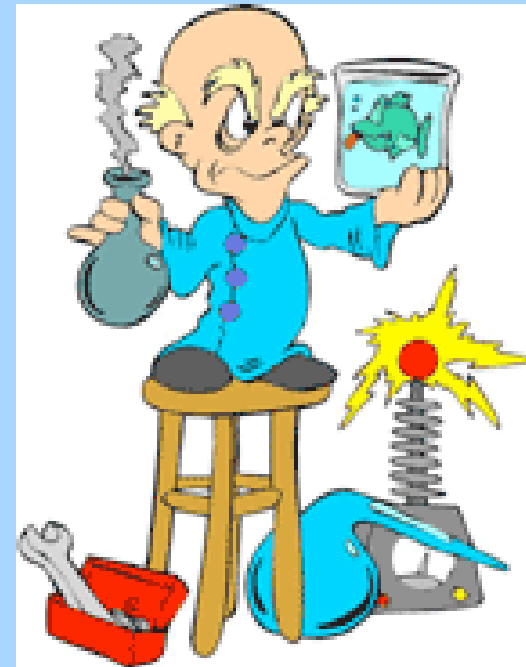
# 1. Introduction

## The Past: Wacky Parallel Machines

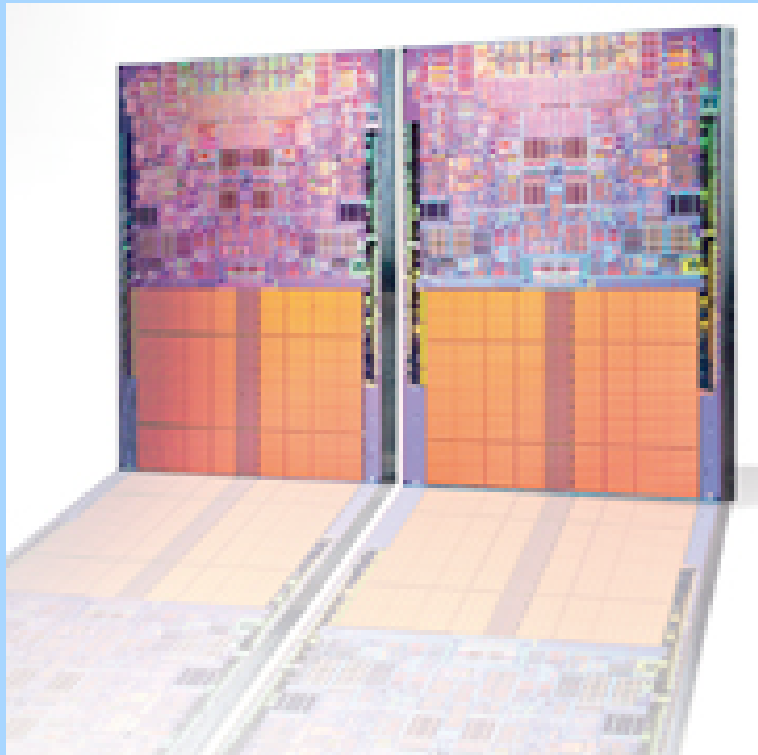


**Illiac IV, 1962 (or 1974?)**

- **Special purpose parallel hardware**
- **Specialist programming notations**
- **Hard to program**
- **Single-processor systems much cheaper and almost as fast**



## The Present: Parallel Programming becomes Mainstream

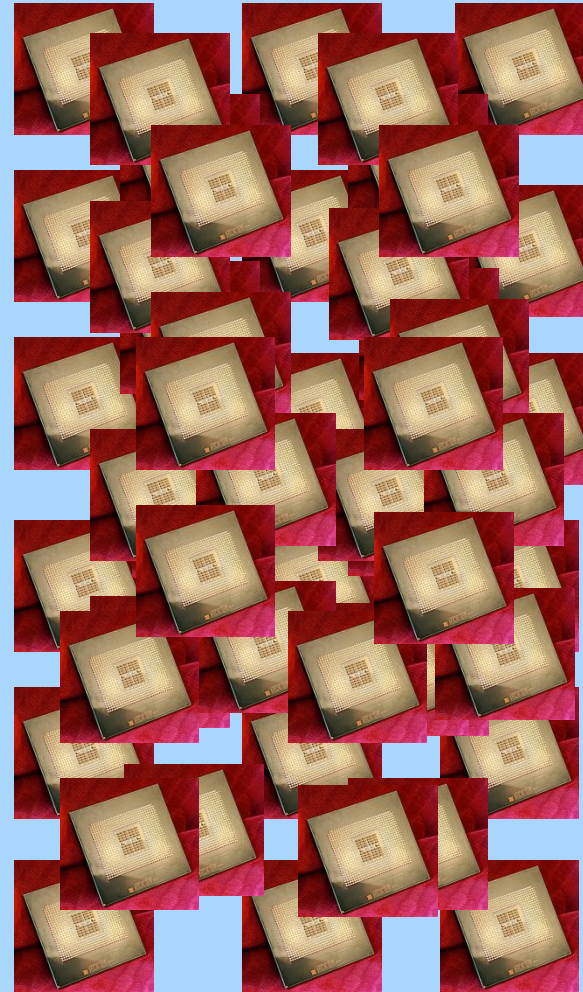


**Intel Xeon 5340, Quad Core, 2007**

- All modern architectures are parallel
- Standard, but low-level programming notations (C+MPI)
- *Still hard to program/exploit*
- Single processor systems now used only in low-end applications

## The Future: Massive Parallelism ...

- **8 Cores**
  - available as twin quad-core
- **16 Cores**
  - available as four quad-core
- **80 Cores?**
  - Intel research project
- **96 Cores?**
  - Clearspeed accelerators
- **1000 Cores?**
  
- **1000000 Cores?**



## ... and the dream of Grid computing

- **Seamless access to computing resources**
  - **Parallel performance available on demand**
  - **easy to use**
  - **no need to configure local systems**
    - **service-based framework**
  - **easy access to databases/other system resources**
  - **pluggable computing power**
- **Wide-area accessibility**
  - **integrates national/international infrastructures**
  - **world-wide availability**



## The Problems for Large-Scale Parallelism

- **Difficult to program parallel computations**
  - **task identification**
  - **placement**
  - **scheduling**
  - **migration**
  - ...
  
- **Symbolic programs can be worse than numeric ones:**
  - **Dynamic data dependencies**
  - **Irregular parallelism (structure and workload)**
  - **Complex symbolic data structures**
  - **Difficult to estimate resource requirements**
  - ...



## Finding the Right Tool

- Users *may not be aware* of existing tools that solve their problems
- Users may not be able to make the *best choice* of tool
- Not realistic to install *all* possible packages locally
  - Maintaining up-to-date working versions
  - Software licensing issues
- Not possible to understand *all* possible packages
  - need to discover operations performed by symbolic packages
- Need a standard taxonomy/semantic interface (à la Frege)
  - different packages have different names for the same operation
  - different packages wrap data differently

## 2. SymGrid



# The SCIENCE Project: Symbolic Computing Infrastructure in Europe



**Five-year, €3.2M EU FP6 Project  
RII3-2005-026133, 2006-2011**

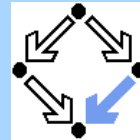
**9 partners:**



**Heriot-Watt (UK), St Andrews (UK), IeAT (Romania),  
RISC (Austria), CNRS (France), Kassel (Germany), TU Berlin (Germany),  
TU/Eindhoven (the Netherlands), MapleSoft (Canada)**



University  
of  
St Andrews



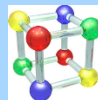
U N I K A S S E L  
V E R S I T Ä T

**4 Systems Initially**

**Maple**



**MuPad**



**Kant**



**GAP**



**"It is reasonable to expect that in the year 2010, the predominant way of doing math will no longer be by pen and paper, but in an integrated web-based math-development system that supports the mathematician in all aspects of mathematics. "**

Michael Kohlhase.

**"And it will be parallel"** Abyd Al Zain



## Overall SymGrid Objectives

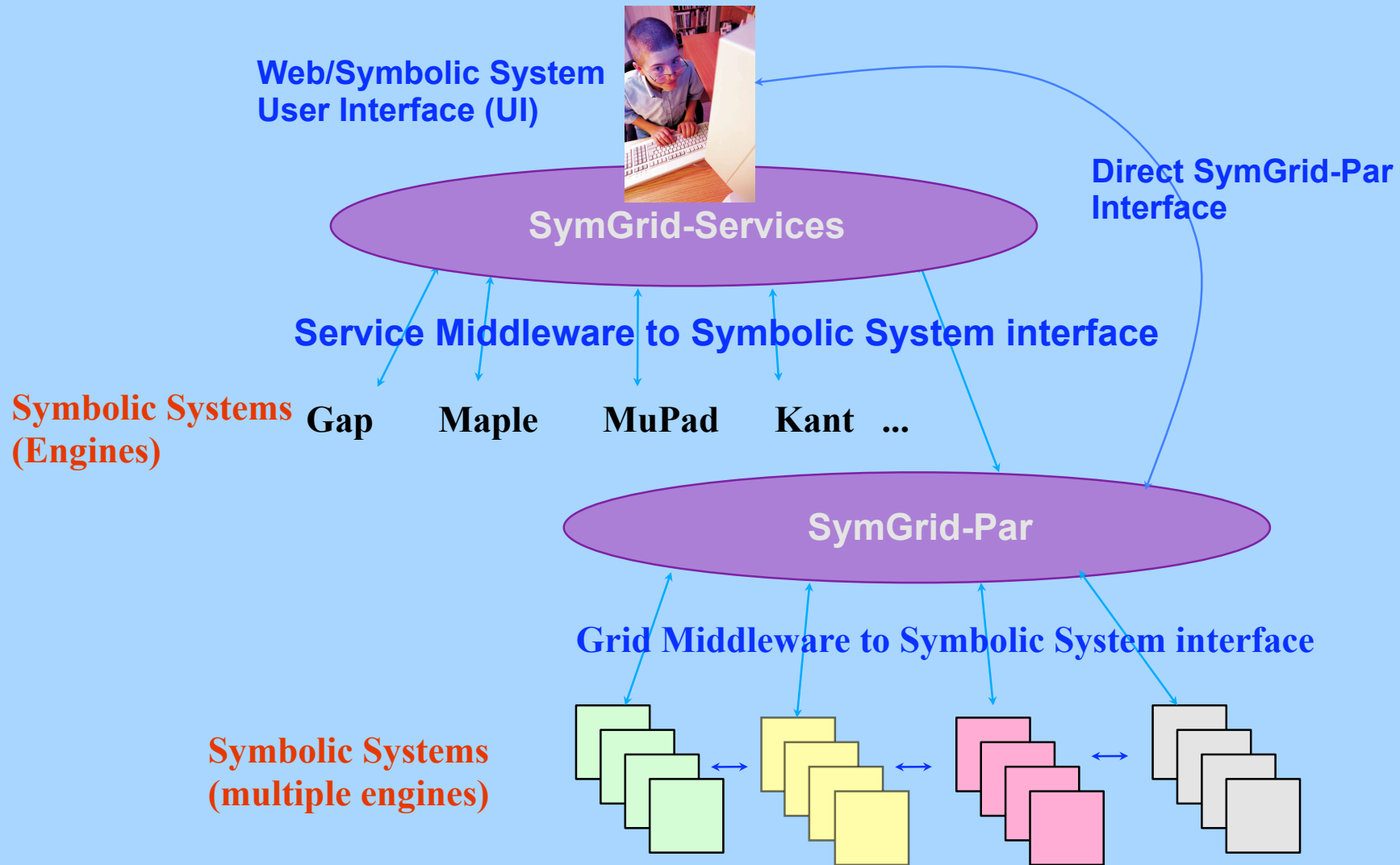
1. produce a portable framework (SymGrid) that will
  - allow symbolic computations **to access** Grid services
  - allow symbolic components **to be exploited** as part of larger Grid service applications on a computational Grid
2. develop resource brokers that will support the irregular workload and computation structures that are frequently found in symbolic computations
3. implement a series of applications that will demonstrate the capabilities and limitations of Grid computing for symbolic computations

**K. Hammond , A. Al Zain , G. Cooperman , D. Petcu , P.W. Trinder:  
SymGrid: a Framework for Symbolic Computation on the Grid.  
EuroPar 2007, Rennes, France, Springer LNCS, August 2007**

## Novel Features

- **Integrate multiple symbolic systems in a single framework**
  - **systems heterogeneity**
  - **platform heterogeneity**
- **Sophisticated integrated *computational steering* interface**
- **New domain-specific patterns of parallel computation**
  - **user-centric view of parallelism**
- **New user community**
  - **may have massive computational demands**
  - **yet exposure to parallelism/Grids is not common**

# SymGrid Design Overview





## Access to Grid Services: SymGrid-Services

- **Middleware to allow generic access to symbolic Grid services.**
- **Service discovery for symbolic Grid components.**
- **Higher-level system interfaces allow straightforward SymGrid access from within symbolic systems. Designed in association with the providers of the Maple, GAP, MuPAD and Kant systems in order to meet the generic requirements of symbolic system providers.**
- **Standard *OpenMath* protocols used to transmit mathematical data (SCSCP)**
  - **XML and native formats**
- **Services may be constructed from heterogeneous components**
- **Services may be *steered* interactively through an integrated user interface**
- ***A. Carstea, M. Frincu, G. Macariu, D. Petcu, K. Hammond: Generic Access to Web and Grid-based Symbolic Computing Services – the SymGrid-Services Framework. International Symp. on Parallel and Distributed Computing, Linz, Austria, IEEE Press, July 2007***



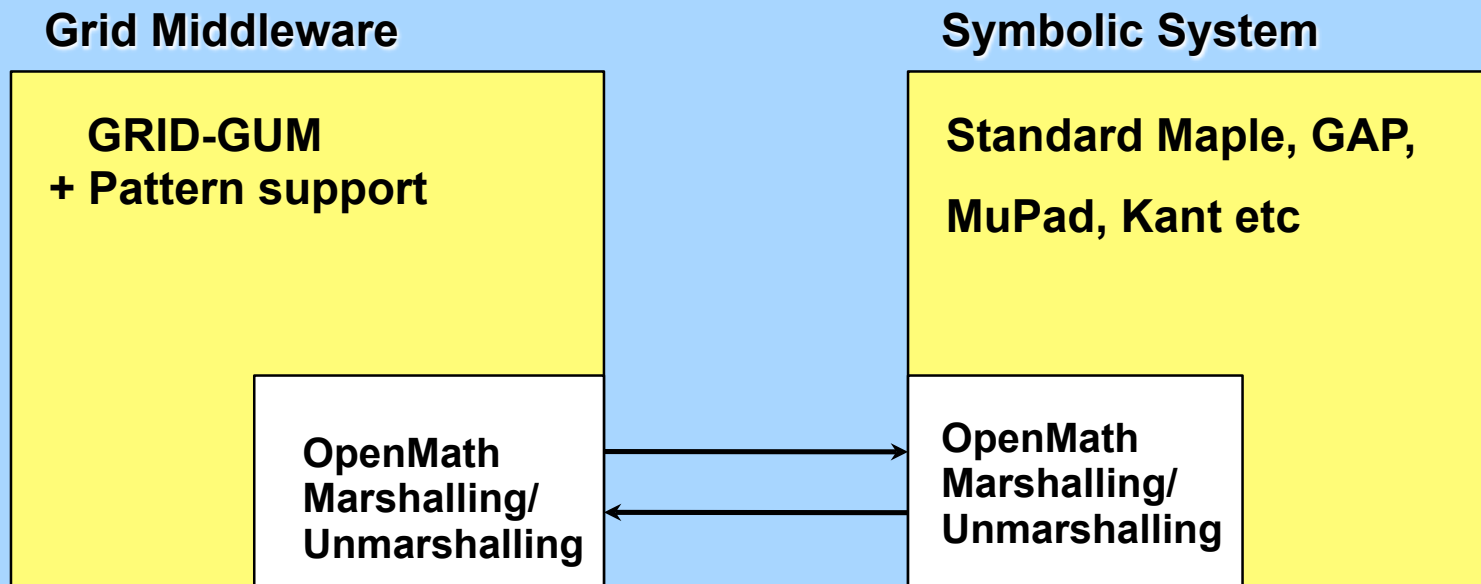


## Symbolic Grid Components: SymGrid-Par

- Orchestrates symbolic components into (possibly parallel) Grid apps.
- Built around GRID-GUM (J. Symb. Comp, 2006)
  - ultra-lightweight thread (*filament*) creation
  - distributed virtual shared memory
  - multi-level scheduling support
  - automatic thread placement
  - automatic datatype-specific marshalling/unmarshalling
  - extends parallel implementation of Haskell
- Links with SymGrid-Services to allow complete Grid-enabled applications to be constructed.

**A. Al Zain, P. Trinder, G. Michaelson, and H-W. Loidl: mGrid-Par: Evaluating a High-Level Parallel Language (GpH) for Computational GRIDs. IEEE Transactions on Parallel and Distributed Systems, 19(2):219--233, 2008.**

## Grid Middleware to Symbolic System Interface

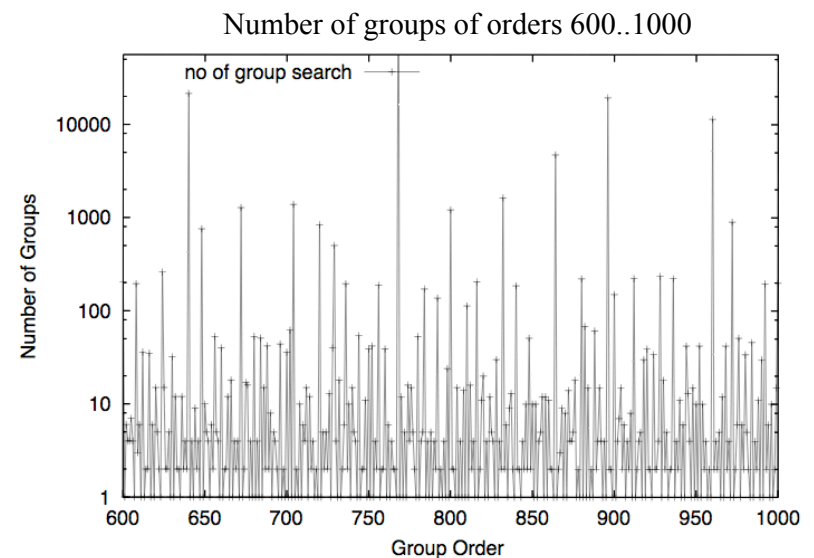


## 3. Results

## Example: smallGroup using the GAP system

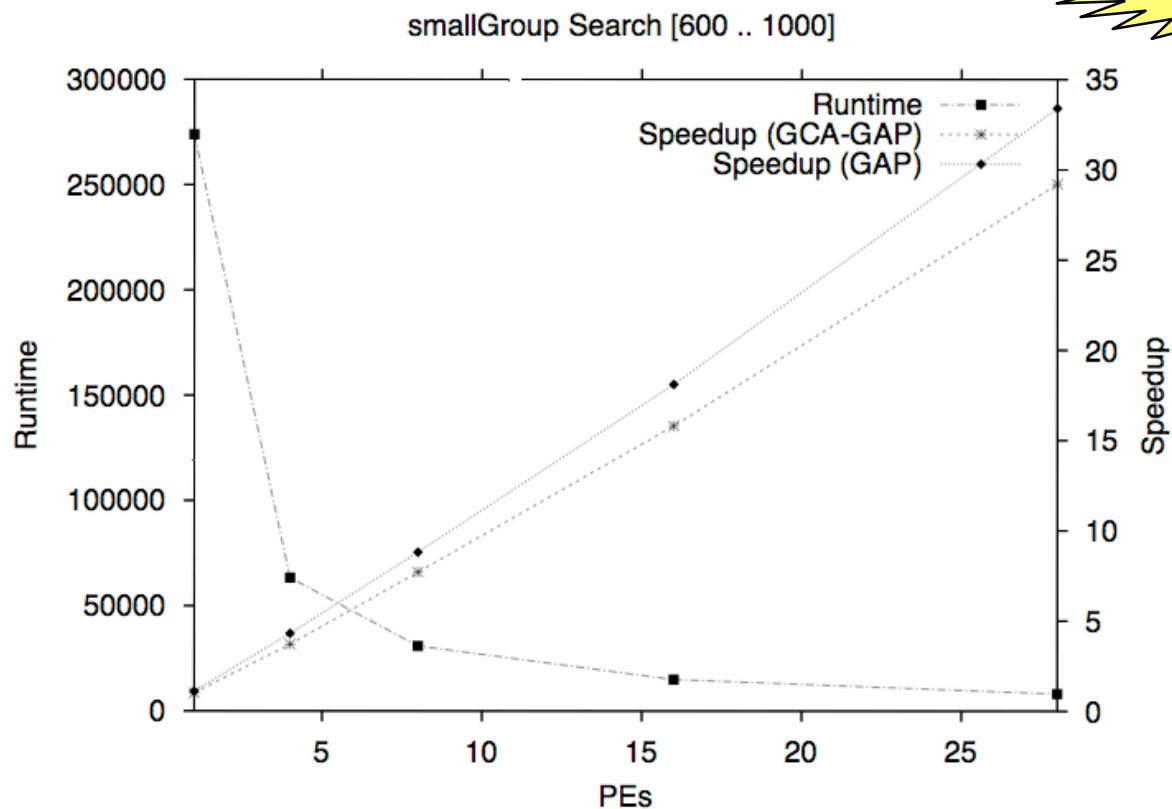
- Search for *groups* with orders within a given range of values, s.t. the average order of the elements is an integer
- Two nested levels of parallelism: search across/within families of groups of the same order
- Two kinds of *irregularity*:
  - *different complexity for different groups*
  - *variation in number of groups of each order*

### Irregularity in Number of Groups



## Single-Cluster Results

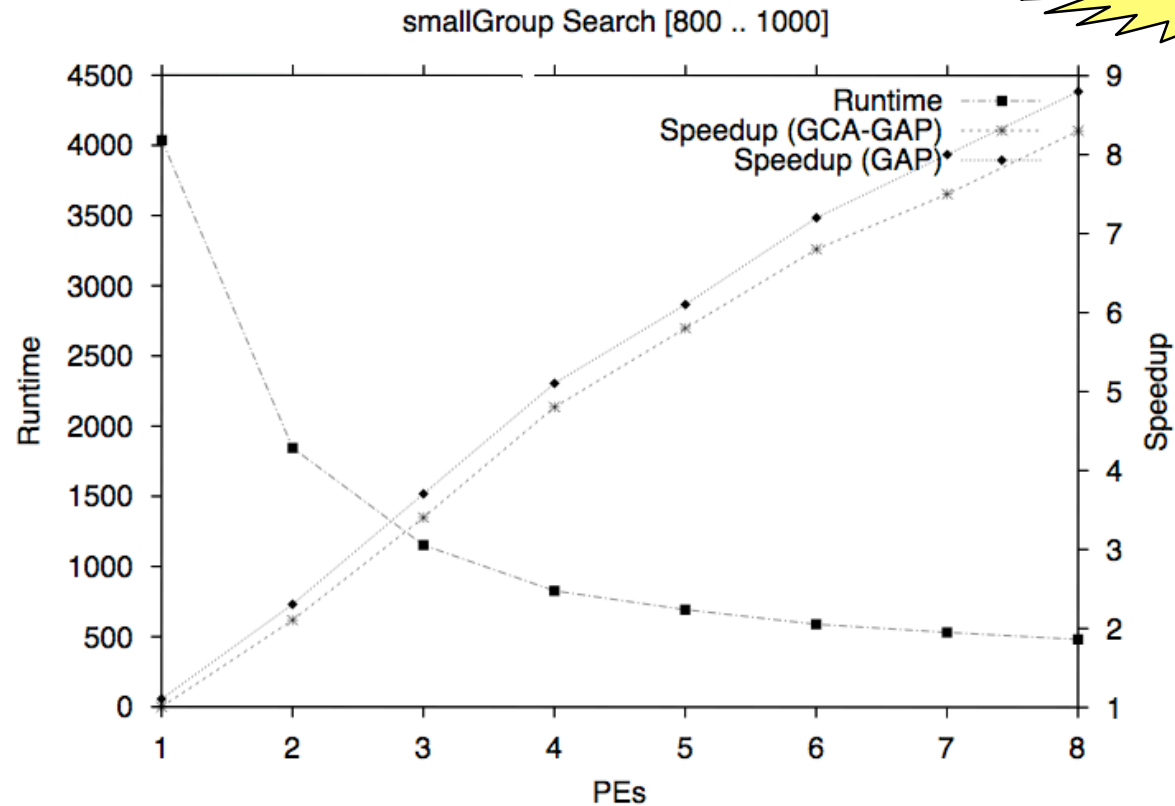
**1,163,006  
candidate groups**



**32 x Intel Pentium IV @ 3.0GHz. 100Mb/s Ethernet network.  
Fedora Linux, kernel version 2.6.10-1.**

## Multicore Results

**42,473  
candidate groups**



**8-core Dell Poweredge 2950, 2 x Intel Xeon 5355 quad-core @ 2.66GHz  
16GB fully-buffered 667MHz DIMMs. Centos Linux 4.5.**

## **4. Composing Grid and Web Services**

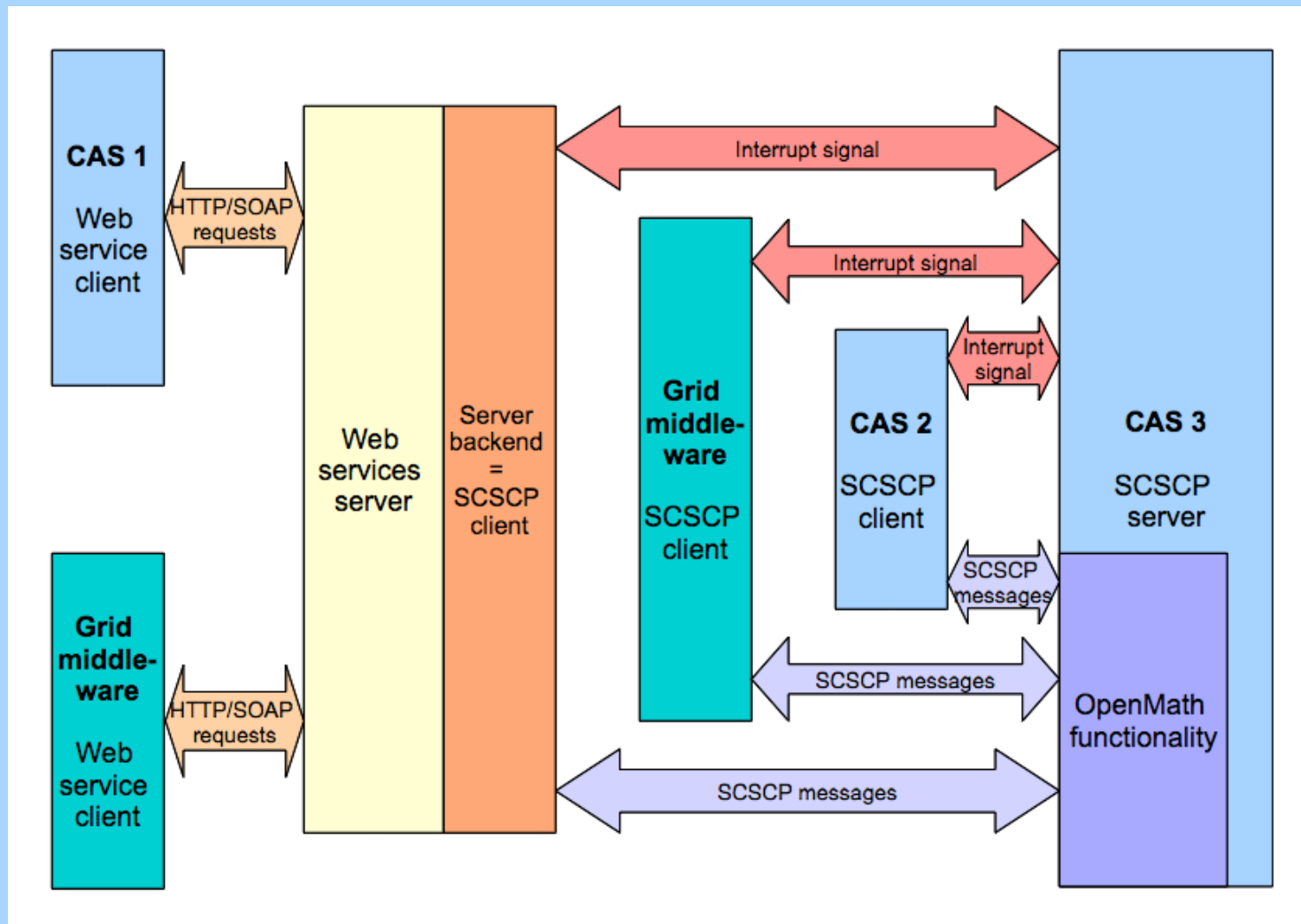


## SCSCP: an OpenMath-based communication protocol

- Using the **Symbolic Computation Software Composability Protocol (SCSCP)**, a computer algebra system (CAS) may offer services for the following clients:
  - i. A Web server which passes on the services as Web services using SOAP/HTTP protocols to other clients
  - ii. Grid middleware (e.g. SymGrid)
  - iii. Another instance of the same CAS
  - iv. Another CAS running on the same computer or remotely



# The BIG Picture: Linking Grid + Web through SCSCP



## 5. Conclusions and Further Work

## Conclusions

- **New framework for symbolic computation**
  - builds on sophisticated, flexible runtime system, **GRID-GUM**
- **Highly generic**
  - two symbolic systems now working: **GAP** and **Maple**
  - other symbolic systems possible (e.g. **Mathematica**, ...)
  - *minimal/no change to underlying sequential system*
- **Supports heterogeneity**
  - systems built from multiple components
- **Promising parallel results on single cluster/multi-core**
- **SCSCP allows linkage to web services**

## Future/Ongoing Work

- **Multi-cluster, geographically distributed**
  - In hand: mainly a matter of practicalities/time
  - Grid clusters in Scotland, Germany, *Romania*
  - Determine whether single-cluster results can be replicated
- **Confirm parallelism results for Maple, ...**
- **Extend to MuPad, Kant, (*your favourite system*)**
  - New UIs for each system
  - New APIs for each system
- **Develop heterogeneous symbolic applications**

## SCIENCE Training schools at RISC, 2007-2010

- The 1st school was held on February 5-18, 2007
- The 2nd school was held on June 25 – July 6, 2007
- One school each year in 2008–2010

### Transnational Access Programme



Supported by the European Commission Framework 6 Programme for Integrated Infrastructures Initiatives under the project SCIENCE



**Transnational Access to RISC**  
RISC-Linz, a research institute at the Johannes Kepler University in Linz, Austria, within the project SCIENCE (2006–2011) offers opportunities to obtain:

- Free access to the infrastructure, facilities, and expertise of a world-leading center in symbolic computation.
- Scientific, technical, administrative, and logistic support, including travel and living expenses.

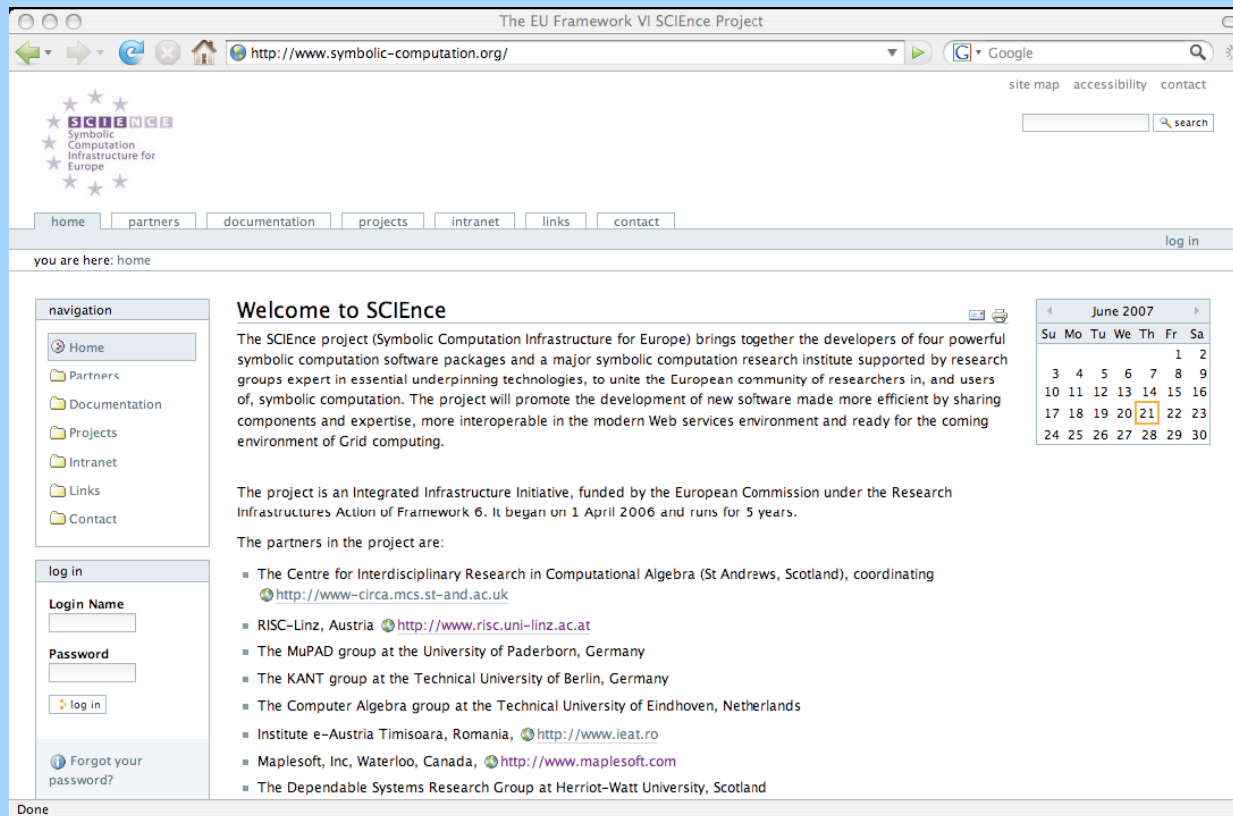
**Targeted Audience**  
Researchers and students interested in using symbolic computation in their work.

**Selection Panel**  
Bruno Buchberger, Arjeh M. Cohen, Marc Giusti, Steve Linton, Peter Paule, Franz Winkler (Chair)

**Scientific Adviser**  
Temur Kutsia

**Detailed Information and Application Procedure**  
<http://www.risc.uni-linz.ac.at/projects/science/access/>

- <http://www.symbolic-computation.org>



The screenshot shows a browser window titled "The EU Framework VI SCIENCE Project" with the URL <http://www.symbolic-computation.org/>. The page features the SCIENCE logo, a search bar, and a navigation menu with links for home, partners, documentation, projects, intranet, links, and contact. A "log in" link is also present.

**you are here: home**

**navigation**

- Home
- Partners
- Documentation
- Projects
- Intranet
- Links
- Contact

**log in**

Login Name:

Password:

[Forgot your password?](#)

**Welcome to SCIENCE**

The SCIENCE project (Symbolic Computation Infrastructure for Europe) brings together the developers of four powerful symbolic computation software packages and a major symbolic computation research institute supported by research groups expert in essential underpinning technologies, to unite the European community of researchers in, and users of, symbolic computation. The project will promote the development of new software made more efficient by sharing components and expertise, more interoperable in the modern Web services environment and ready for the coming environment of Grid computing.

The project is an Integrated Infrastructure Initiative, funded by the European Commission under the Research Infrastructures Action of Framework 6. It began on 1 April 2006 and runs for 5 years.

The partners in the project are:

- The Centre for Interdisciplinary Research in Computational Algebra (St Andrews, Scotland), coordinating  
<http://www-circa.mcs.st-and.ac.uk>
- RISC-Linz, Austria <http://www.risc.uni-linz.ac.at>
- The MuPAD group at the University of Paderborn, Germany
- The KANT group at the Technical University of Berlin, Germany
- The Computer Algebra group at the Technical University of Eindhoven, Netherlands
- Institute e-Austria Timisoara, Romania, <http://www.ieat.ro>
- Maplesoft, Inc, Waterloo, Canada, <http://www.maplesoft.com>
- The Dependable Systems Research Group at Herriot-Watt University, Scotland

Calendar for June 2007:

Su	Mo	Tu	We	Th	Fr	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

# *Questions*