Distributed Symbolic Computations

Dana Petcu

Western University of Timisoara
And Institute e-Austria Timisoara, Romania

petcu@info.uvt.ro

Overview

- The problem
- The past
- The present
- The future?

The problem

Symbolic computation?

- Automation of the steps of mathematical problem solving
- Modern field of CS and Math

  which deals with symbolic objects, e.g. logical or algebraic formulae, rules or programs
- Main goal is exactness
- Examples of operations: diff, integral, factors, roots

Subfields of symbolic computing

- ***computer algebra (CA)***,
- automated theorem proving,
- computational combinatorics,
- computational geometry,
- automated programming,
- functional or logic programming.

Symbolic methods - applications

- computer aided design
- software development
- VLSI design
- geometric modelling
- reasoning
- robot programming
- human genome

etc

Problems behind CA systems

Lagging relative to numerical computing,

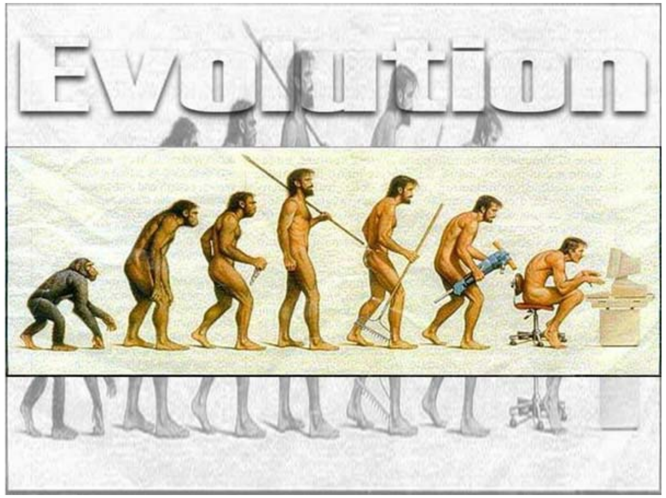
mainly due to the inadequacy of available computational resources:

1. computer memory
2. processor power.

Solution: parallel & ***distributed CA***

- Solving larger problems
- Build new algorithms
- Build new systems

The past



Distributed computations emerge

From: Tilmann Bubeck, Martin Hiller, Wolfgang Küchlin, and Wolfgang Rosenstiel. Distributed Symbolic Computation with DTS. In *Proc. of Parallel Algorithms for Irregularly Structured Problems*, LNCS, Lyon, France, Sep 1995. Springer. to appear.

Distributed Symbolic Computation with DTS

T. Bubeck* M. Hiller W. Küchlin** W. Rosenstiel

W. Schickard Institute for Informatics
University of Tübingen
Sand 13, D-72076 Tübingen, Germany
EMail: {bubeck,hiller,kuechlin,rosen}@informatik.uni-tuebingen.de
`http://www-ti.informatik.uni-tuebingen.de/dts`

Keywords: Parallel and Distributed Computing, Threads, Algebraic and Symbolic Computing, Irregular Load Balancing.

Abstract. We describe the design and implementation of the *Distributed Threads System* (DTS), a programming environment for the parallelization of irregular and highly data-dependent algorithms. DTS extends the support for *fork/join* parallel programming from shared memory threads to a distributed memory environment. It is currently implemented on top of PVM, adding an asynchronous RPC abstraction and turning the net into a pool of anonymous compute servers. Each node of DTS is multithreaded using the C threads interface and is thus ready to run on a multiprocessor workstation. We give performance results for a parallel implementation of the RSA cryptosystem, parallel long integer multiplication, and parallel multi-variate polynomial resultant computation.

General reasons for parallelism

- ability to reduce the wall-clock time i.e. the user’s waiting time for the solution

  problems that are *processor bound*

- Ability to solve problems that cannot fit into memory of a ``workstation''

  problems that are *memory bound*

Memory bound is the driving force in parallelization of CA algorithms

Parallel symbolic algs. – 1990-2000

- Fast multiprecision integer arithmetic, eg. factorization, large integer multiplication
- Fast polynomial arithmetic, eg. GCD, factorization, differentiation
- Fast exact solution of linear equation systems
- Fast solution of polynomial equation systems via Groebner basis

M.Matooane, Parallel systems in symbolic and algebraic computation, PhD Thesis, 2001, University of Cambridge

Difficulties in building parallel or distributed symbolic computing systems

- Unpredictable data dependencies,
- Algorithmic dependence on irregular data which are difficult to be dynamically partitioned,
- In case of distributed memory, one processor can exceed its available memory, while there is still space available globally
- Complexity of some of algebraic computations limiting ability to estimate resource requirements
- developing completely new systems
  - is efficient,
  - but difficult
  - usually only a few parallel algorithms within such a system are fully implemented and tested

Adding parallelism

- develop CASs for shared memory architecture
- develop computer algebra hardware;
- add parallel primitives for communication and cooperation to existing CASs;
- build distributed memory systems based on standard communication middleware;
- build distributed systems for loosely coupled machines or across the Internet.

Example: Maple

Author	***Tool***	***Technology***
Char,'90	Sugarbush	C/Linda + Maple
Wang,'91	Parallel	Shared files + Maple
Siegl,'93	||Maple||	Strand + Maple
Bernardin,'97,	Maple for Paragon	Rewrite Maple
Diaz,Kartofen,'98	FoxBox	MPI + Maple
Schreiner,'98,	Distributed Maple	Java + Maple
Petcu, '01,	PVMaple	PVM + Maple

Schreiner et al., Distributed Maple-parallel computer algebra in networked environments J.Symbolic Comput.35(3),2003

Mathematical software

Thousands of codes of all kinds performing all kinds of symbolic computations:

- ***General purpose systems:***

  Axiom Aldor Derive Macsyma Magma Maple Mathematica MuPAD Reduce etc

- ***Special purpose systems:***

  ACE Albert Algeb Amore Bergman Cannes / Parcan Carat Casa Chevie C-Meataxe CoCoA Crep Desir Discreta Felix Fermat FoxBox Gap GiNaC Kan/sm1 Kant LiDIA LiE BIGLIE Macaulay Mas Masyca Moc NTL Pari Parsac Quotpic ReDuX RepTiles SAC-1 Aldes/SAC-2 Saclib SciNapse Senac Simath Singular SymbMath Symmetrica Theorema Theorist etc

- ***Packages:***

  Arep Cali CLN Crack, LiePDE ApplySym ConLaw Dimsym EinS FeynArts FormCalc FeynCalc Grape Molgen Orme Ratappr TTC etc

- Johannes Grabmeier, Erich Kaltofen, Volker Weispfenning , Computer Algebra Handbook, Springer, 2003

- CAS: http://www.symbolicnet.org

User needs

- Users *may not be aware* of existing tools to solve their math problems
- Users may not be able to make *best choice*
- *Not realistic to install* all packages locally
  - – Know specifics of all software
  - – Maintain up-to-date licenses of all software
  - – Even for rarely used ones
- Need to normalize, categorize, and discover operations performed by mathematical packages
- Need for a standard taxonomy
  - different packages perform the same operation under different names.
- Need semantic interface to abstract packages peculiarities

Challenge: find ways to re-use codes

- Standard representation of mathematical objects

  solved recently

- Standard way to invoke, compose and discover mathematical packages

  in progress

Mathematical Representation: OpenMath

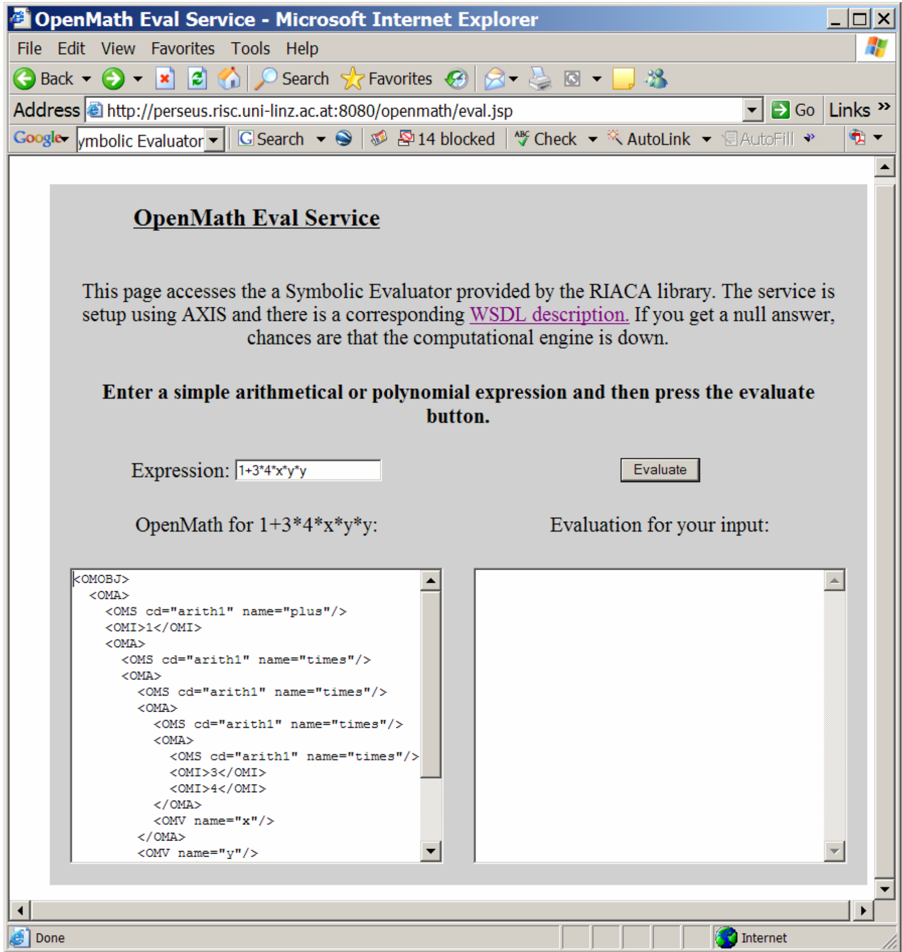
- Data formats for portable mathematical objects: OpenMath, MathML, OMDoc
- OpenMath: Standard developed by an European research consortium.
  - Abstract syntax model for mathematical objects
    - variable, symbol
    - quantier (variable, object), application(object, object), annotation(object, object)
  - Concrete syntax representations
  - Content dictionaries (CDs)
    - Collections of constant (function/predicate) symbols
    - Standard set of CDs plus extensions

OpenMath

Maple: int(sin(x),x=1..10);

OpenMath

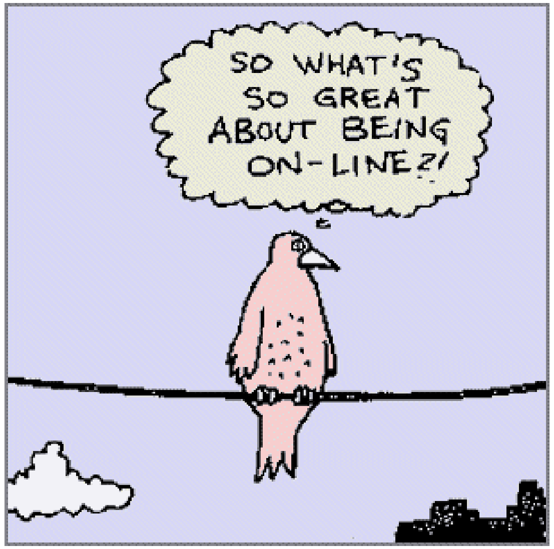
```
<OMOBJ>
  <OMA>
    <OMS cd="calulus1" name="defint"/>
    <OMA>
      <OMS cd="interval1" name="interval/>
      <OMI> 1 </OMI>
     <OMI> 10 </OMI>
    </OMA>
    <OMBIND>
      <OMBVAR>
        <OMV name="x"/>
      </OMBVAR>
      <OMA>
        <OMS cd="transc1" name="sin"/>
        <OMV name="x"/>
      </OMA>
    </OMBIND>
  </OMA>
</OMOBJ>
```



Protocols and APIs for communications

- IAMC, MathWeb, JavaMath etc
- IAMC: Internet-Accessible Mathematical Computation
  - HTTP-like protocol for server-client communication:
  - Informal description of service provided.
  - Abstract protocol for service access
  - Requires insight to be used
- MathWeb
  - Software bus combining mathematical services
  - Broker providing access object for service by name
  - Abstraction from service locations and from object encodings

The present

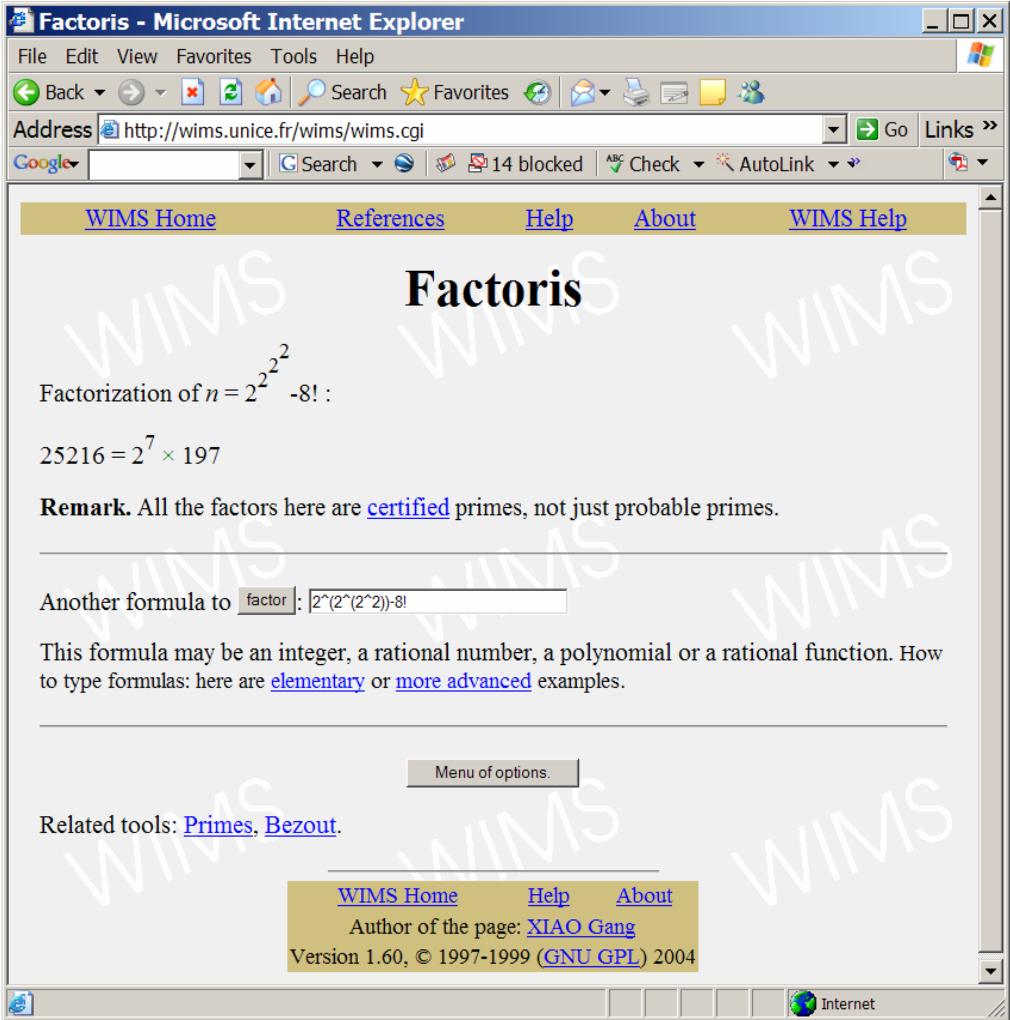


E-Mathematics

- Resources:
  - bibliographic data (lots, with meta-data)
  - papers (lots, HTML)
  - ***software (some, user interfaces)***
- Services:
  - citation indexes (very used)
  - ***computations (seldom used)***

Web-enabled systems

- MapleNet and webMathematica
- WIMS: WWW Interactive Mathematics Server

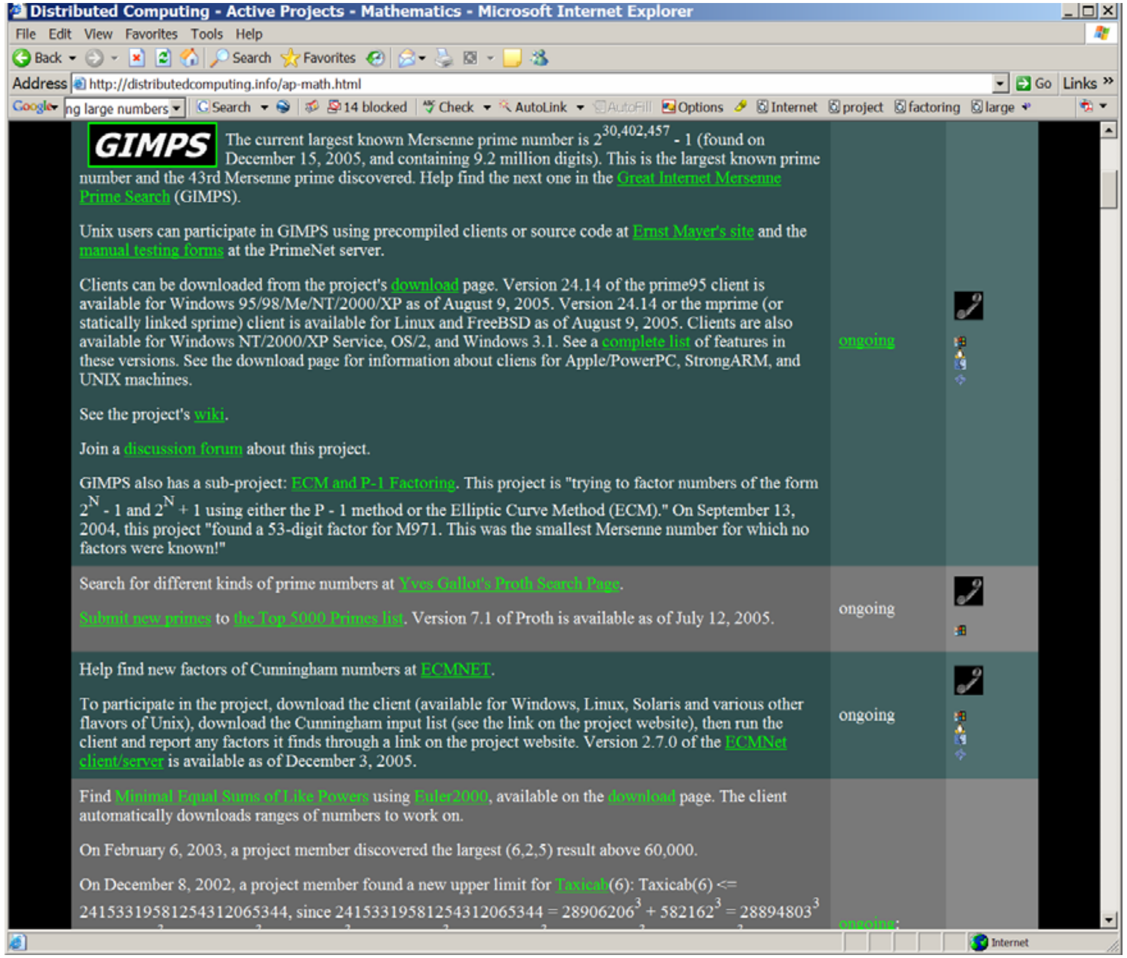


Internet projects

http://distributedcomputing.info/ap-math.html

- finding large prime numbers,
- factoring large numbers,
- computing digits of Pi,
- finding collisions on known encryption algorithms

etc.



Towards automatization

Consumers of resources and clients of services:

- Humans
- Databases
- ***Software***
- ***Other resources and services***

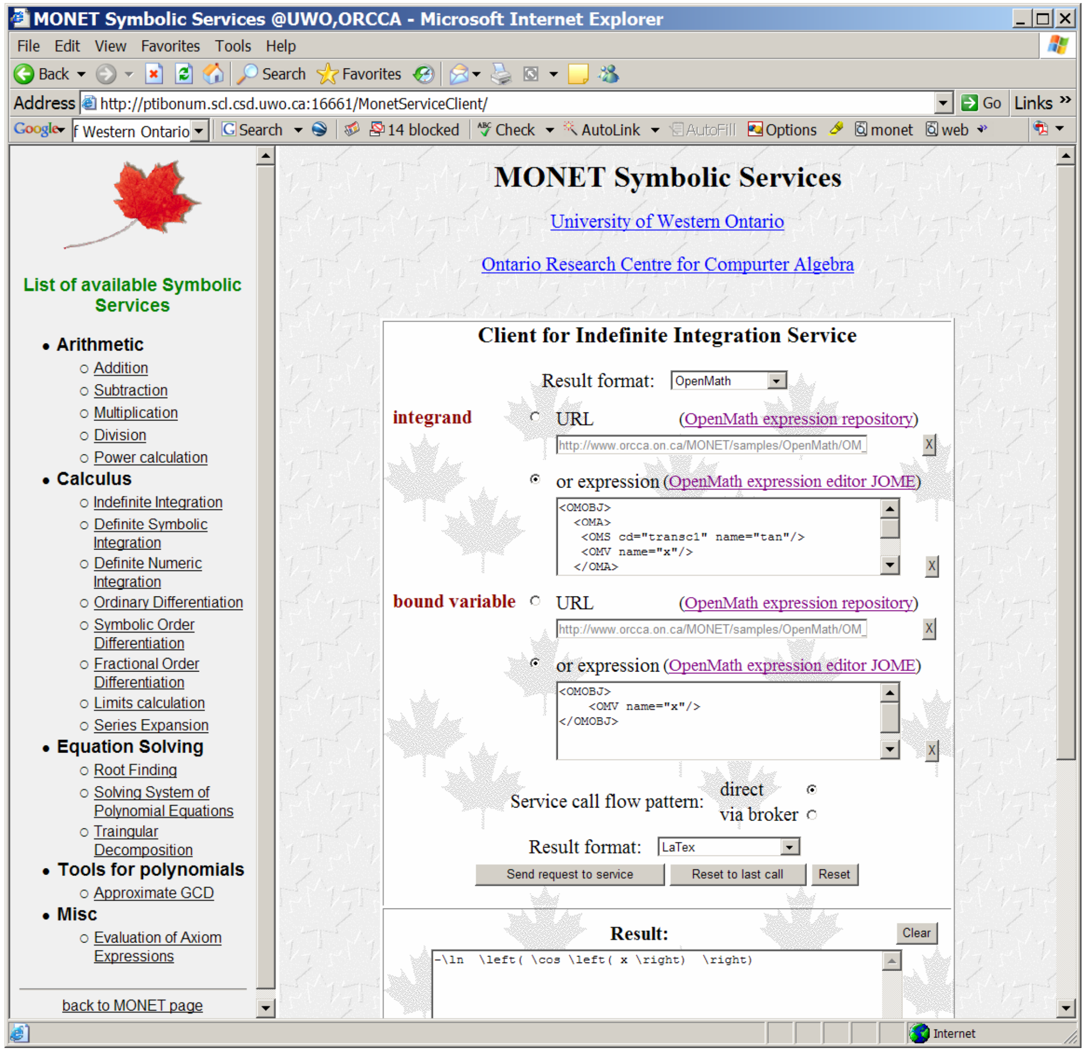
Initiatives for Mathematical Web services

- MONET (2002-2004) demonstrated the applicability of the semantic Web to the world of mathematical software (discover services dynamically)
- MathWeb-SB (2003) – access via broker by name
- MathBroker (2005-2007) – Web registry to publish/discover

MONET

- Ability to *discover services* dynamically based on published descriptions which express both their mathematical and non-mathematical attributes

- A *symbolic solver wrapper* was designed



Initiatives for Mathematical Grid services

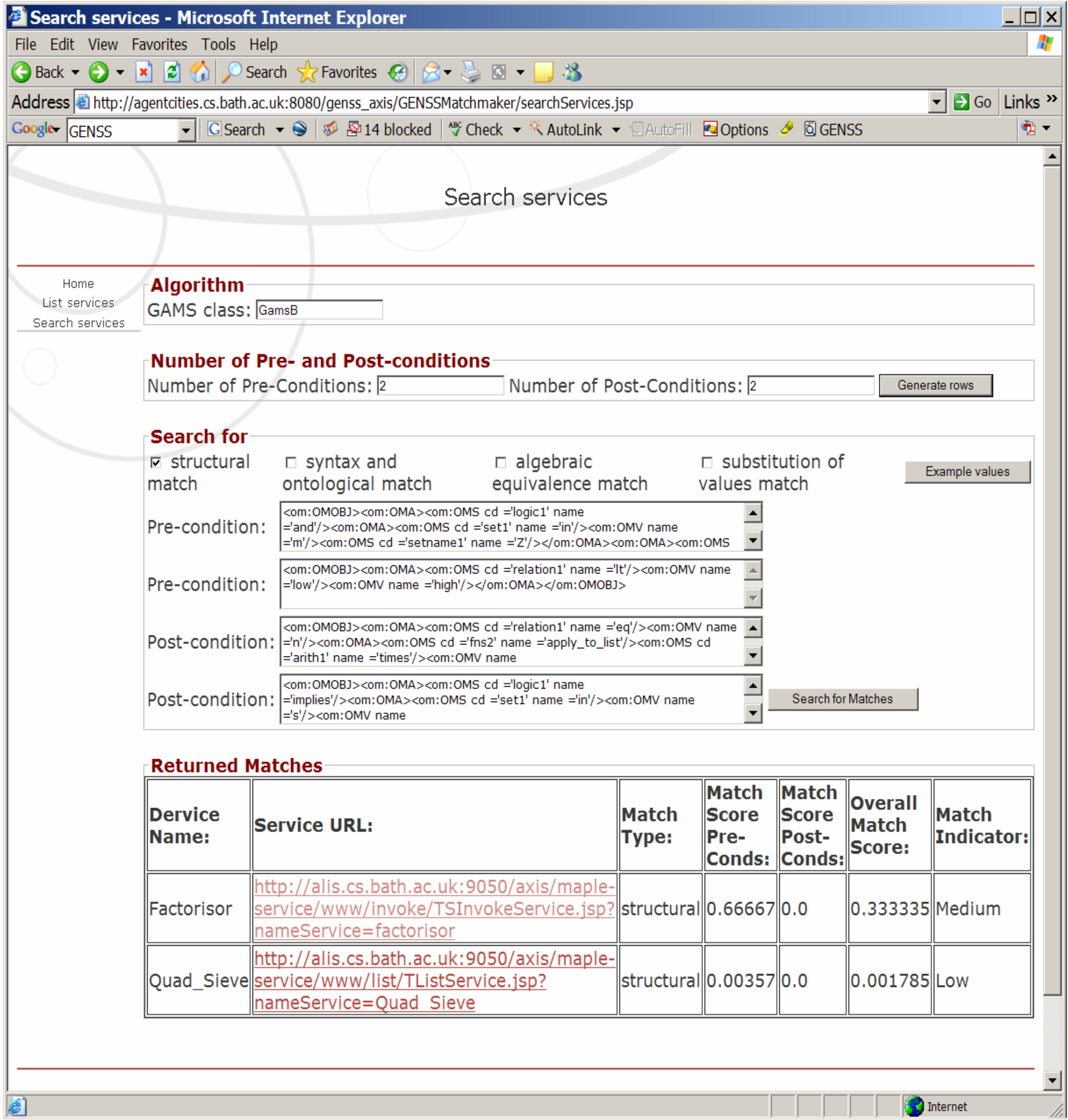
- Why Grid?
  - Whereas the Web is a service for sharing information over the Internet, the Grid is a **service for sharing computer power and data storage** capacity over the Internet
  - High potential as discovery accelerator
  - Way to categorize, explore, discover, invoke and compose thousand of software packages
- GENSS – follows MONET, research on advertisement and discovery, ontology
- GEMLCA – deploy a legacy code

D.Petcu, D.Tepeneu, M.Paprzycki, T.Ida, Symbolic Computations on Grids, Chapter 6 in the book "Engineering the Grid: status and perspective", eds.

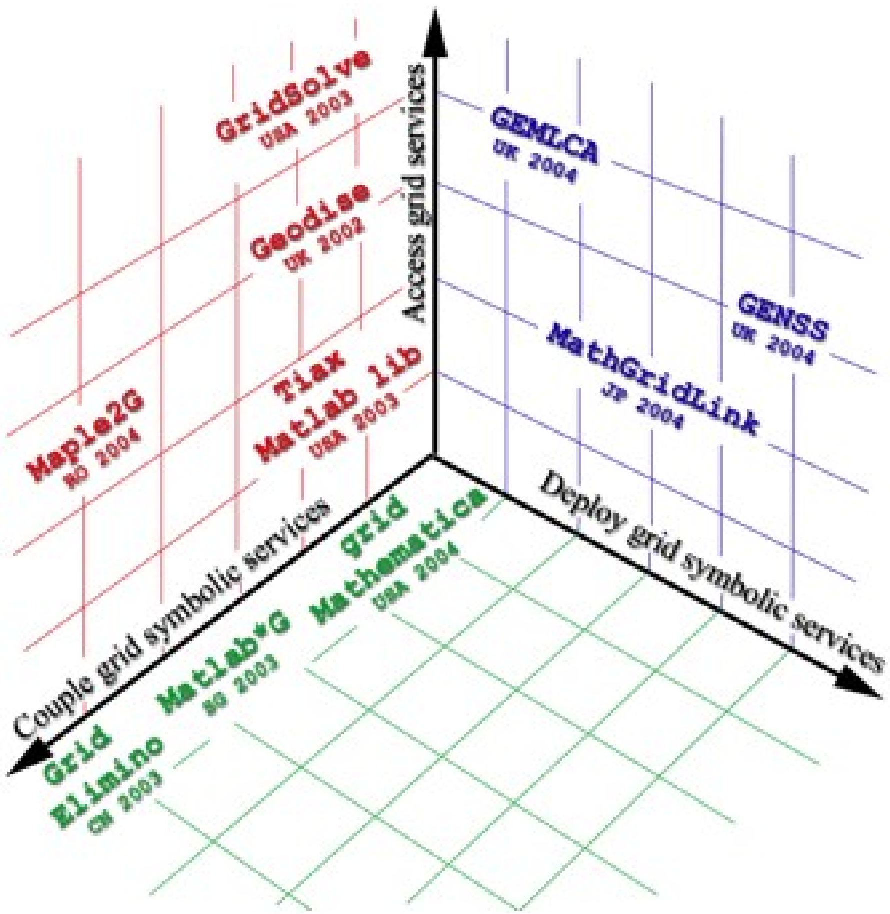
Beniamino di Martino, et al ASP, 2006, pp. 91-107

GENSS

- Grid Enabled Numerical and Symbolic Services, initiated in 2004
  http://genss.cs.bath.ac.uk/index.htm
- research was focused in two areas:
  1. matchmaking techniques for advertisement and discovery of mathematical services,
  2. design and implementation of an ontology for symbolic problems



Symbolic computations on Grids



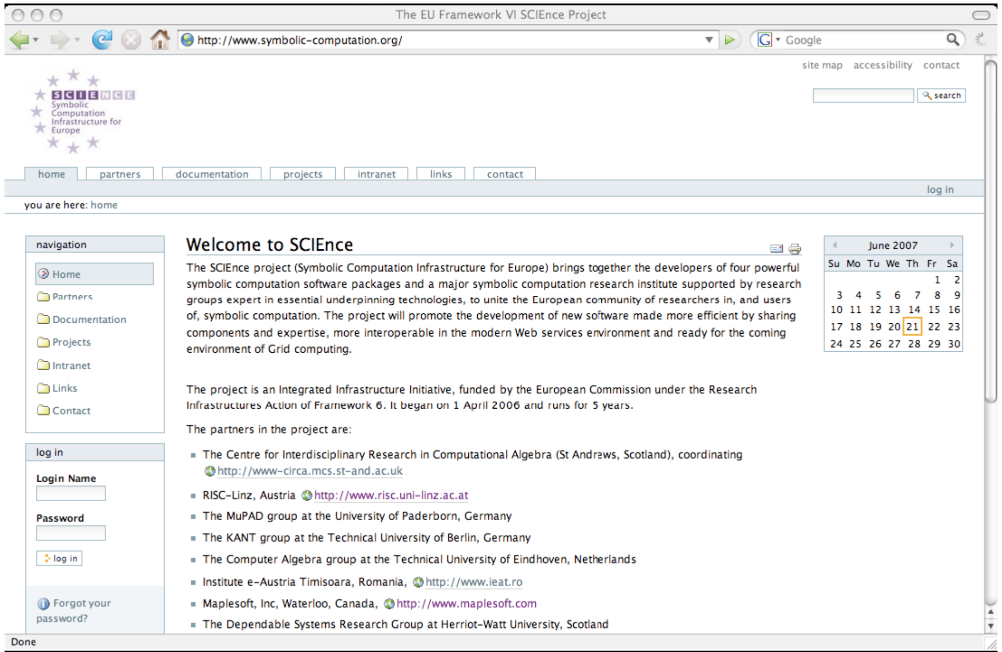
SCIEnce project (2006-2011)

SIXTH FRAMEWORK PROGRAMME

SCIEnce – **S**ymbolic **C**omputation **I**nfrastructure in **E**urope

http://www.symbolic-computation.org

5-year project supported by the EU Framework VI programme grant RII3-CT-2005-026133.



R&D in SCIEnce

- Software composability:
  - a programme of standards developments and implementations for symbolic computation software to use Web services and OpenMath technologies, allowing them to be efficiently composed to solve complex problems
- Symbolic computing on Grids:
  - developing common standards and middleware to allow the production of Grid-enabled symbolic computation systems constructing research prototypes supporting appropriate security, scheduling, and resource broking for complex symbolic computing applications on computational Grids

SCIEnce consortium

- University of St Andrews, St Andrews, UK
- Research Institute for Symbolic Computation, Linz, Austria
- Centre National de la Recherche Scientifique, France
- Universität Paderborn, Germany
- Technische Universiteit Eindhoven, Netherlands
- Technische Universität Berlin, Germany
- Institute e-Austria Timisoara, Romania
- Maplesoft, Waterloo, Canada
- Heriot Watt University, Edinburgh, UK

- **GAP**
- **KANT/KASH**
- **Maple**
- **MuPAD**

SCIEnce: Software composability

Designed the Symbolic Computation Software Composibility Protocol (**SCSCP**) :

by which a computer algebra system (CAS) may offer services for the following clients:

- a Web server which passes on the same services as Web services using SOAP/HTTP protocols to another clients
- Grid services
- Another instance of the same CAS (in a parallel computing context)
- Another CAS running on the same computer or remotely

A. Konovalov, S. Linton. Symbolic Computation Software Composability Protocol Specification. CIRCA preprint 2007/5, University of St Andrews. http://www-circa.mcs.st-and.ac.uk/preprints.html

SCIEnce: using OpenMath

- All messages in the protocol are represented as OpenMath objects, using the new Content Dictionary **cascall1** developed for this purpose.
- **SCSCP** specifies: Semantical and technical descriptions of OpenMath-encoded messages to and from CAS:
  - remote procedure call
  - returning result of successfully completed procedure
  - returning a signal about procedure termination

  and also allowed sequences of these messages.

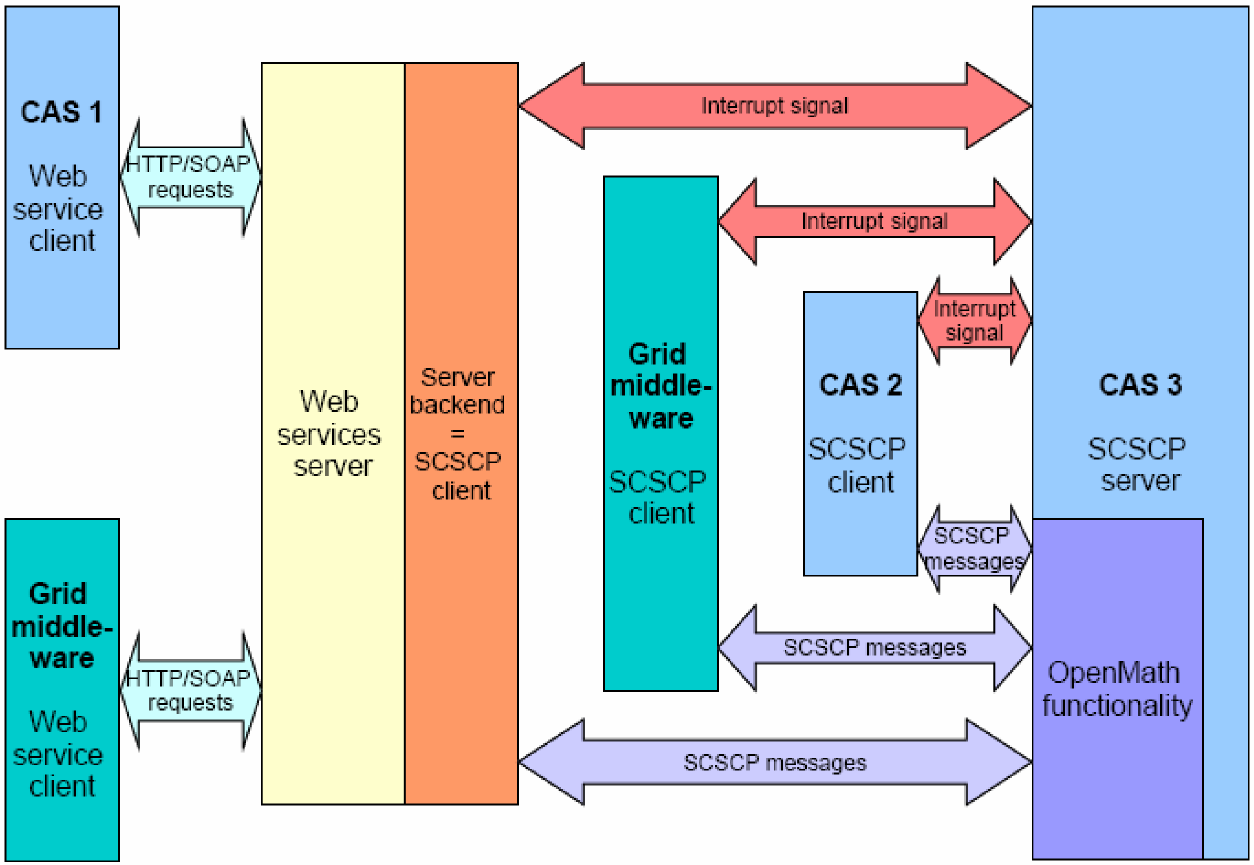
D. Roozemond. OpenMath Content Dictionary cascall1. http://www.win.tue.nl/SCIEnce/cds/cascall1.html

SCIEnce: cascall1

In the next example we retrieve the group [24,12] from GAP Small Groups Library, creating it at the GAP side and requesting a cookie for it (omitted options will be set to default values):

```
<OMOBJ>
    <OMATTR>
        <OMATP>
            <OMS cd="cascall1" name="call_ID" />
            <OMSTR>alexk_9053</OMSTR>
            <OMS cd="cascall1" name="option_return_cookie" />
            <OMSTR/>
        </OMATP>
        <OMA>
            <OMS cd="cascall1" name="procedure_call" />
            <OMSTR>GroupByCatalogueNumber</OMSTR>
            <OMI> 24</OMI> <!-- Argument 1 -->
            <OMI> 12</OMI> <!-- Argument 2 -->
        </OMA>
    </OMATTR>
</OMOBJ>
```

SCIEnce: Current vision of SCSCP



SCIEnce: Work in progress

- Adding basic SCSCP implementations to all systems
- Identifying and developing new OpenMath content dictionaries and other standards extensions needed
- Adding support of selected OpenMath CDs in all systems
- Implementing higher level interfaces in all systems

Overall objectives for symbolics on Grids

- Produce a portable framework (**SymGrid-Services**) that will allow:
  - symbolic computations to access Grid services
  - symbolic components to be exploited as part of larger Grid service applications on a computational Grid
- Develop resource brokers that will support the irregular workload and computation structures that are frequently found in symbolic computations (**SymGrid-Par**)
- Implement a series of applications that will demonstrate the capabilities and limitations of Grid computing for symbolic computations

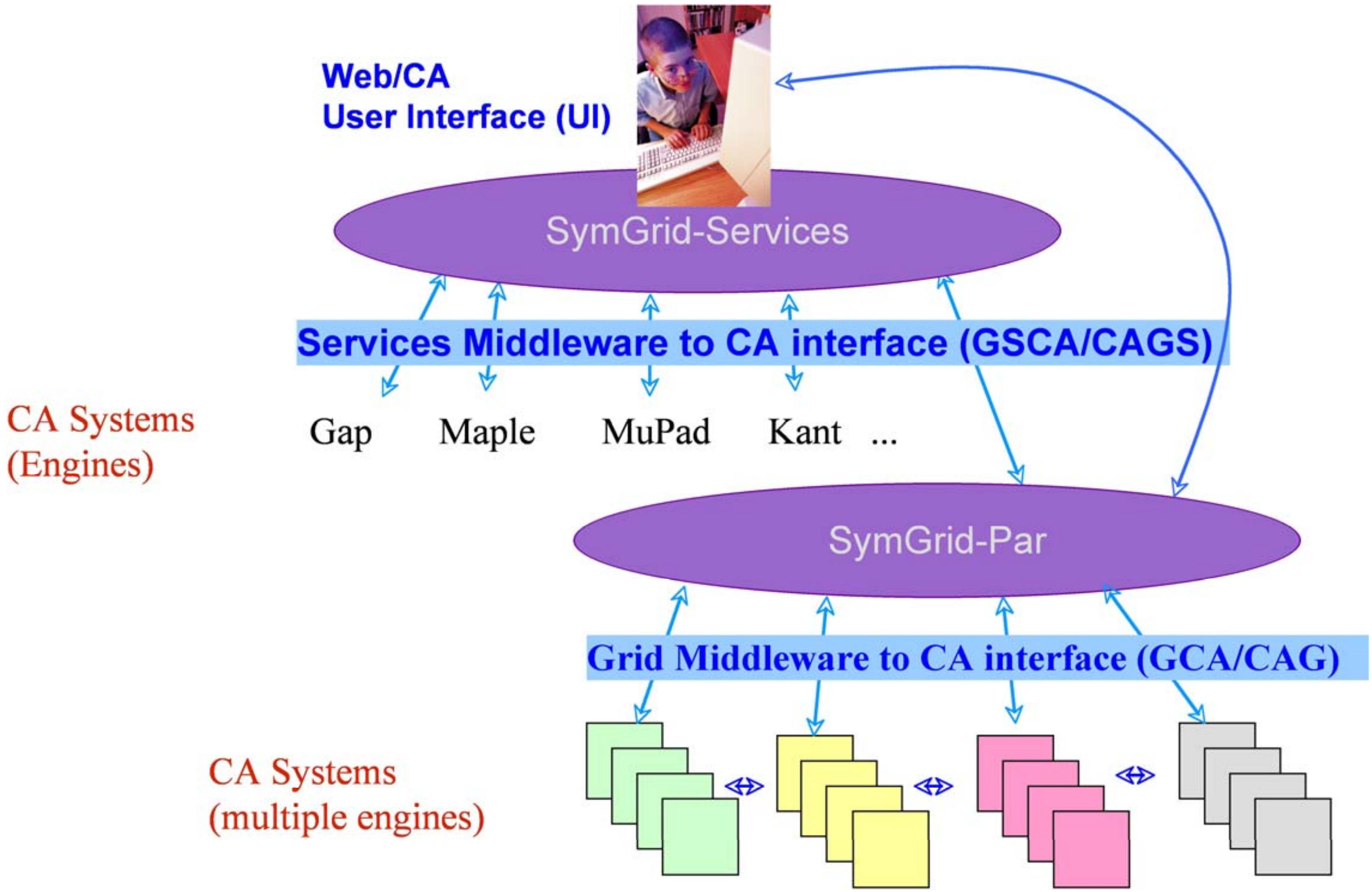
SCIEnce's SymGrid

- A new middleware is developed – **SymGrid** – to allow the construction of large-scale Grid-enabled symbolic applications.
- SymGrid components:
  - **SymGrid-Services** to provide access to Grid services from symbolic applications
  - **SymGrid-Par** to support the construction of high-performance applications on computational Grids.
- **SymGrid-Par** is built around **GRID-GUM**, a system designed to support parallel computations on Grids and adapted to interface with symbolic computation engines

Al Zain, P. Trinder, H.-W. Loidl and G. Michaelson. Managing Heterogenity in a Grid Parallel Haskell. J. Scalable Comp.: Practice and Experience, 6(4), 2006.

A. Carstea, M. Frincu, G. Macariu, D. Petcu, K. Hammond. Generic Access to Web and Grid-based Symbolic Computing Services – the SymGrid-Services Framework. Proc. ISPDC 07

SymGrid architecture



Training schools at RISC, 2007-2010

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

PASCO

Symbolic-Numeric Computation 2007
www.orcca.on.ca/conferences/snc2007

July 25-27, 2007

Parallel Symbolic Computation 2007
www.orcca.on.ca/conferences/pasco2007

July 27-28, 2007

Joint Speakers
A D Kennedy, U Edinburgh
K O Geddes, U Waterloo (after dinner)

SNC 2007

Invited Speakers

André Galligo, U Nice
Erich Kaltofen, NCSU
Nick Trefethen, U Oxford
Charles Wampler, GM Research
Lihong Zhi, MMRC CAS

Topics

- Hybrid symbolic-numeric algorithms
- Approximate polynomial GCD and factorization
- Symbolic-numeric methods for polynomial systems
- Structured matrices in symbolic-numeric computation
- Differential equations for symbolic-numeric computation
- Symbolic-numeric algorithms for algebraic geometry, geometric computation and optimization
- Implementation of symbolic-numeric algorithms
- Model construction with approximate algebraic algorithms
- Applications of symbolic-numeric computation
- Numerical algebraic geometry

Important Dates

Submission deadline:	April 9, 2007
Notification:	May 28, 2007
Camera ready version due:	June 15, 2007

Program Committee

Dario Bini, Italy
Robert Corless, Canada
James Demmel, USA
Ioannis Emiris, Greece
Marc Moreno Maza, Canada
Bernard Mourrain, France
Victor Pan, USA
Greg Reid, Canada
Tateaki Sasaki, Japan
Andrew Sommese, USA
Jan Verschelde (Chair), USA
Dongming Wang, France
Zhonggang Zeng, USA

PASCO 2007

Invited Speakers

Mike Bauer, UWO
Matteo Frigo, Cilk Arts
Thierry Gautier, INRIA
Katherine Yelick, UC Berkeley

Topics

- Parallel computer algebra
- High performance for exact and approximate procedures
- Analysis of parallel algorithms for algebraic computations
- Parallel computing for number theory, combinatorial and discrete methods
- Distributed data-structures for algebraic computation
- Implementations of solvers on multi-cores, SMPs, clusters, supercomputers and grids
- Interactive parallel symbolic computation
- Volunteer computing for symbolic problems
- Applications of parallel symbolic computation

Important Dates

Submission deadline:	April 16, 2007
Notification:	May 28, 2007
Camera ready version due:	June 15, 2007

Program Committee

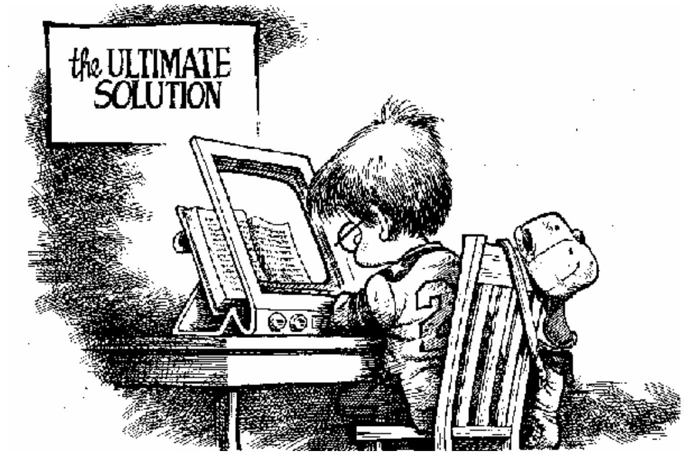
Gene Cooperman, USA
Jean-Guillaume Dumas, France
Jean-Charles Faugère, France
Mark Giesbrecht, Canada
Erich Kaltofen, USA
Anton Leykin, USA
Marc Moreno Maza (Chair), Canada
Jean-Louis Roch, France
David Saunders, USA
William Stein, USA
Carlo Traverso, Italy
Gilles Villard, France

General Chairs: Stephen M. Watt (SNC) watt@orcca.on.ca; Marc Moreno Maza (PASCO) moreno@orcca.on.ca
Administration: Meg Borthwick meg@csd.uwo.ca
Local Arrangements: Oleg Golubitsky oleg@orcca.on.ca; Éric Schost schost@orcca.on.ca
Publicity: François Lemaire lemaire@lifl.fr

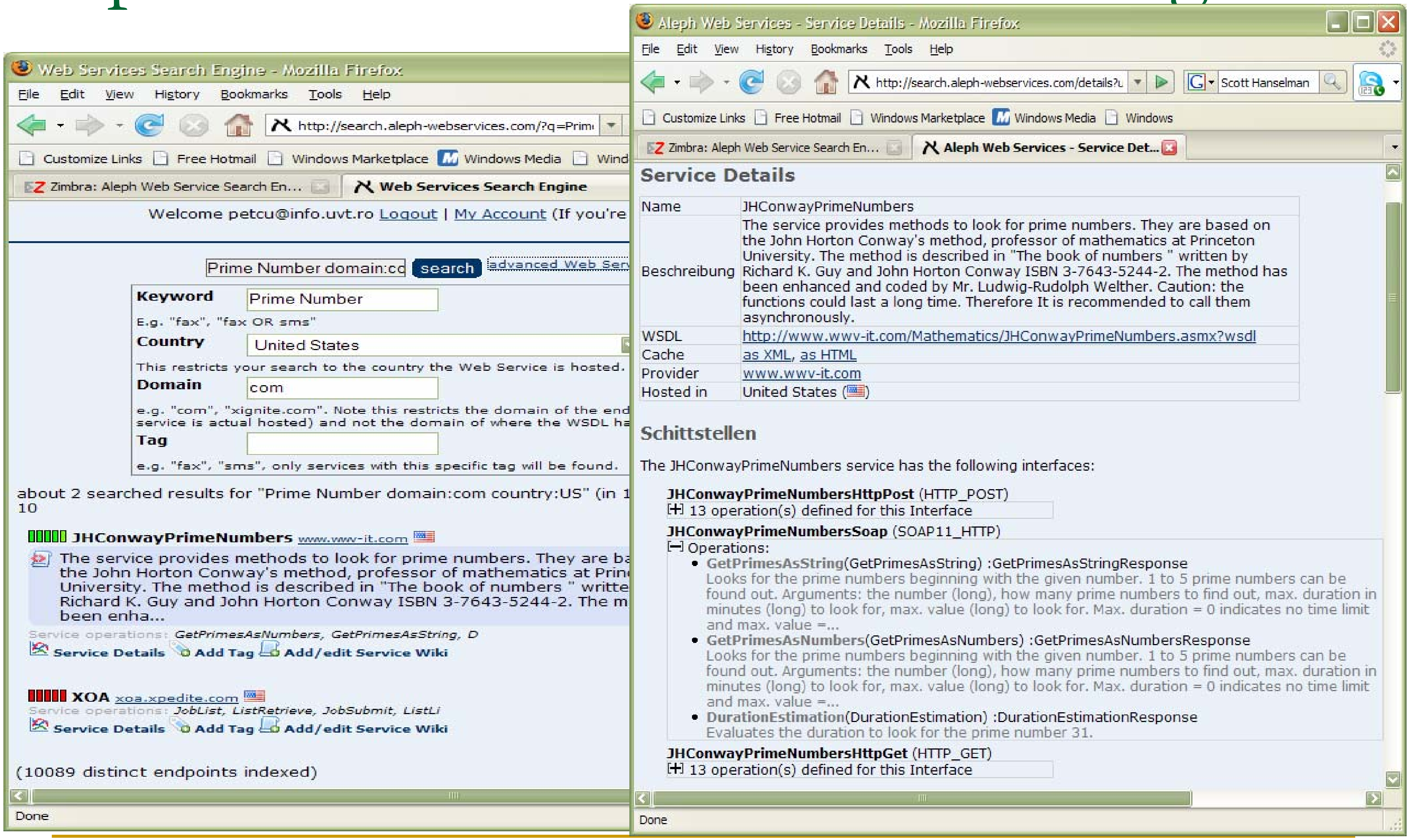
MITACS NSERC CRSNG ORCCA FIELDS Western acm In Cooperation Maplesoft SHARCNET

University of Western Ontario, London Canada

The future

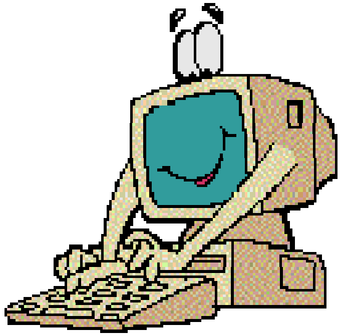


Specialized services and search engines



Dynamic service composition

The current trend is to
create environments with self-
combining mathematical services



“The market of Five computers”

n's Computer Zen - The Skynet Compute Cloud: I think there is a world market for maybe Five Computers - Mozilla Firefox

listory Bookmarks Tools Help

http://www.hanselman.com/blog/TheSkynetComputeCloudIThinkThereIsAWorldMarketForMaybeFiveComputers.aspx

Free Hotmail | Windows Marketplace | Windows Media | Windows

Scott Hanselman's Computer Ze... | Scott Hanselman: Value of Writing Win...

> *"I think there is a world market for maybe five computers"*

but it might have been **Thomas Watson** in 1943, president of IBM, but it might be a myth.

The thinking was that a number of large mainframey computers would be built and used by the world. That of course, didn't happen, as my watch has more memory than **ENIAC** did, and there's computers everywhere, but as the web continues to mature, I believe that things will **conflate** in the next 10-15 years and as more and more companies begin to

- outsource storage to services like **S3**
- move servers into **Virtual Machines**
- move all email and apps to **Google Apps**

All this is building a collective trust with large entities like Amazon, and as prices fall with uptimes rising, more companies will say, "*who am I to build a datacenter? I'll just host in one of The Five.*"

Here's a list of services (by no means exhaustive) from Wikipedia's **Utility Computing article**, as of today:

- **Amazon S3** - Bulk storage and bandwidth for static content
- **Amazon EC2** - Pay by the hour CPU
- **NearlyFreeSpeech** - Pay as you go web hosting for web pages, dynamic content, domains, DNS, etc
- Sun Microsystems **Sun Grid** - Pay by the CPU hour
- **Strikeiron** Web Services Marketplace - Pay per Web API call.
- **USi** an **AT&T** company - USiPinnacle - Pay-as-you-go enterprise applications
- **Zimki** - J avascript based utility computing system. Pay by bandwidth, storage, and javascript operations.
- **ElasticLive** Utility web hosting service based on Amazon's Elastic Compute Cloud
- **InsynQ** utility computing services
- **CPAASP** online accounting solutions for on-demand enterprises
- **Distributed Potential** Pay-Per-Use Grid Computing Capacity
- **Dell Inc.** have a specific Cloud Computing Solution through their Datacenter Solutions Division

I believe this list will likely turn into **The Five Computers**:

1. Google (Apps)
2. Amazon (EC2)
3. Microsoft (Live)
4. Sun (Grid)
5. A network of installed bots on every small computer, possibly built into the OS, to use idle CPU cycles for the collective.

Expect to solve them

Calculus and Analysis
Discrete Mathematics
Foundations of Mathematics
Geometry
History and Terminology
Number Theory
Probability and Statistics
Recreational Mathematics
Topology

Alphabetical Index
Interactive Entries
Random Entry
New in *MathWorld*

MathWorld Classroom

About *MathWorld*
Contribute an Entry
Send a Message to the Team

Order book from Amazon

Last updated:
12,615 entries
Sun Aug 20 2006

Created, developed, and nurtured by Eric Weisstein at Wolfram Research

Unsolved Problems

COMMENT On this Page

There are many unsolved problems in mathematics. Some prominent outstanding unsolved problems (as well as some which are not necessarily so well known) include

1. The Goldbach conjecture.

2. The Riemann hypothesis.

3. The conjecture that there exists a Hadamard matrix for every positive multiple of 4.

4. The twin prime conjecture (i.e., the conjecture that there are an infinite number of twin primes).

5. Determination of whether NP-problems are actually P-problems.

6. The Collatz problem.

7. Proof that the 196-algorithm does not terminate when applied to the number 196.

8. Proof that 10 is a solitary number.

9. Finding a formula for the probability that two elements chosen at random generate the symmetric group S_n.

10. Solving the happy end problem for arbitrary n.

11. Finding an Euler brick whose space diagonal is also an integer.

12. Proving which numbers can be represented as a sum of three or four (positive or negative) cubic numbers.

13. Lehmer's Mahler measure problem and Lehmer's totient problem on the existence of composite numbers n such that $\phi(n) \mid (n-1)$, where $\phi(n)$ is the totient function.

14. Determining if the Euler-Mascheroni constant is irrational.

15. Deriving an analytic form for the square site percolation threshold.

16. Determining if any odd perfect numbers exist.

The Clay Mathematics Institute (`http://www.claymath.org/millennium/`) of Cambridge, Massachusetts (CMI) has named seven "Millennium Prize Problems," selected by focusing on important classic questions in mathematics that have resisted solution over the years. A $7 million prize fund has been established for the solution to these problems, with $1 million allocated to each. The problems consist of the Riemann hypothesis, Poincaré conjecture, Hodge conjecture, Swinnerton-Dyer Conjecture, solution of the Navier-Stokes equations, formulation of Yang-Mills theory, and determination of whether NP-problems are actually P-problems.

In 1900, David Hilbert proposed a list of 23 outstanding problems in mathematics (Hilbert's problems, a number of which have now been solved, but some of which remain open. In 1912, Landau proposed four simply stated problems, now known as Landau's problems, which continue to defy attack even today. One hundred years after Hilbert, Smale (2000) proposed a list of 18 outstanding problems.

K. S. Brown, D. Eppstein, S. Finch, and C. Kimberling maintain webpages of unsolved problems in mathematics. Classic texts on unsolved problems in various areas of mathematics are Croft *et al.* (1991), in geometry, and Guy (1994), in number theory.

SEE ALSO: Beal's Conjecture, Catalan's Conjecture, Fermat's Last Theorem, Hilbert's Problems, Kepler Conjecture, Landau's Problems, Mathematics Contests, Mathematics Prizes, Poincaré Conjecture, Problem, Solved Problems, Szemerédi's Theorem, Twin Primes. [Pages Linking Here]

Instead conclusions

"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 sys.
that supports the mathematician
in all aspects of mathematics. "

Michael Kohlhase,
MathWeb project (http://www.mathweb.org/)

QUESTIONS ?