



Accelerating Scientific Discovery

through powerful
collaborative platforms

Marie-Christine Sawley

General Manager CSCS

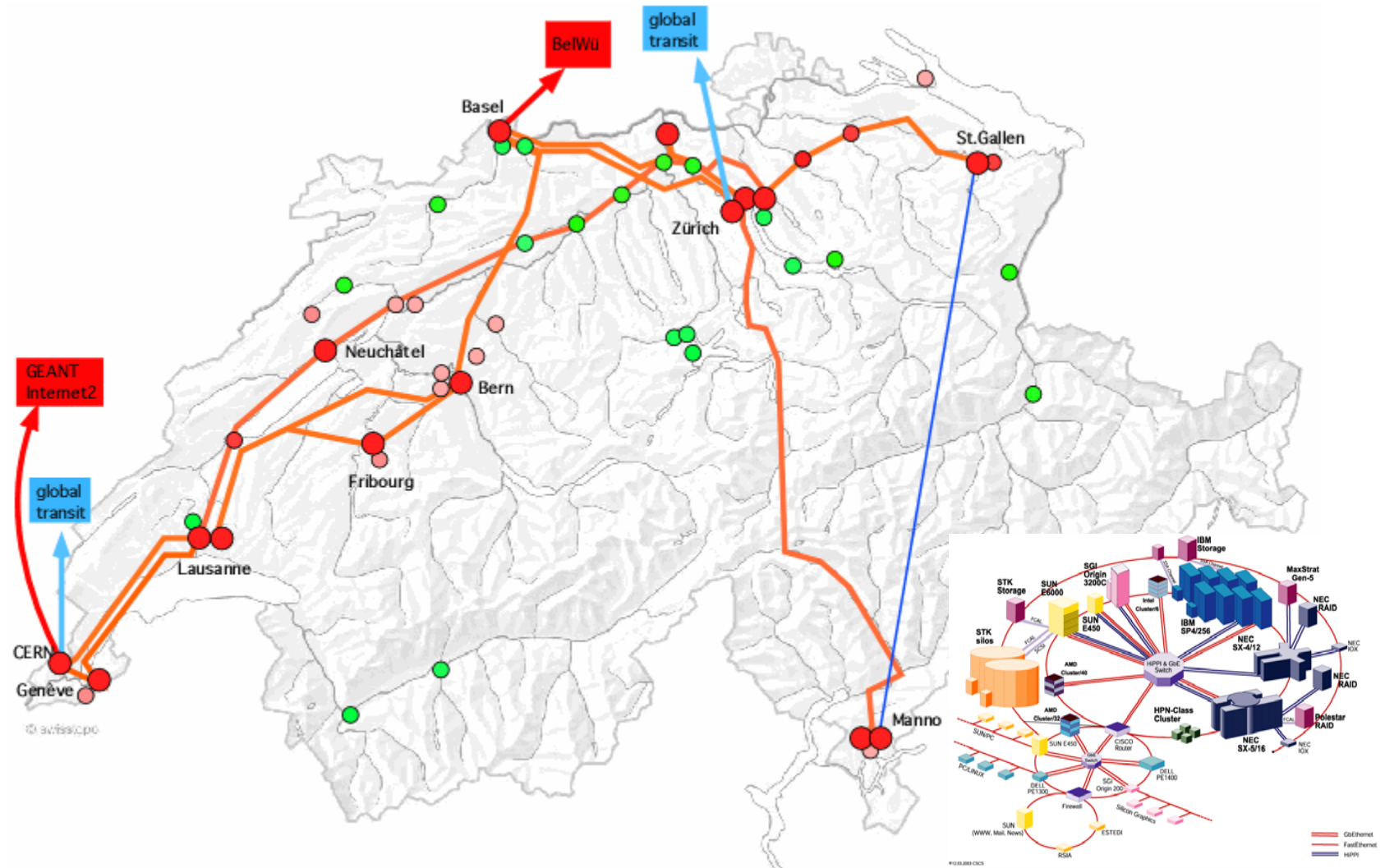
Cern, 25 March 2003

Outline of the presentation

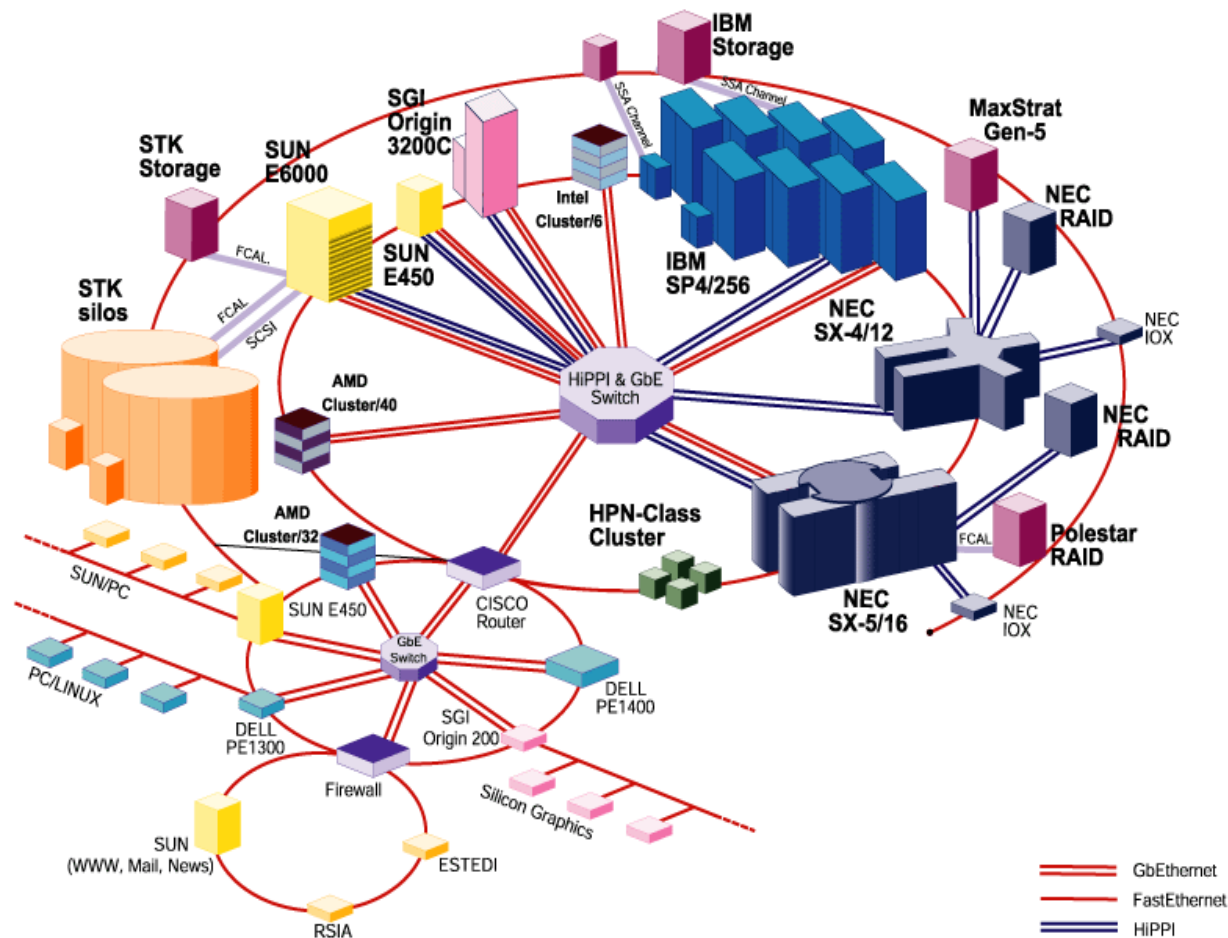


- Brief recap on 2003
- Strategy 2004-2007
- CSCS autonomy and new governance
- First developments

HB networks make locality less of an issue



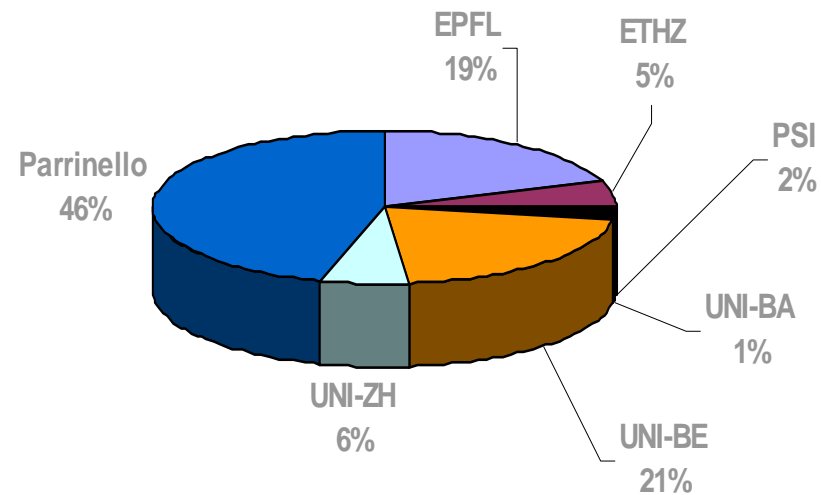
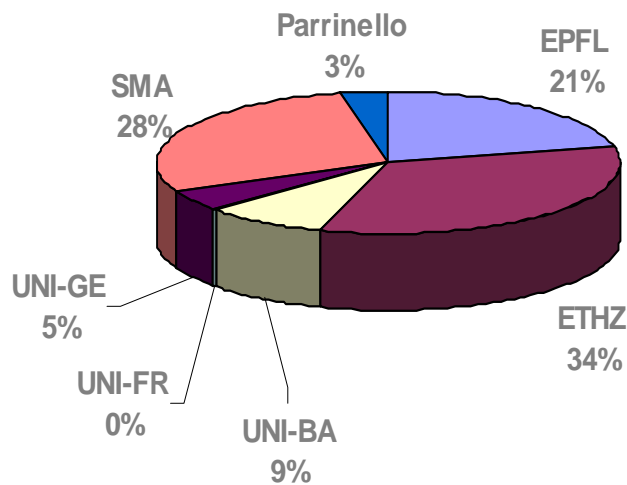
CSCS configuration



Usage by institution 2003

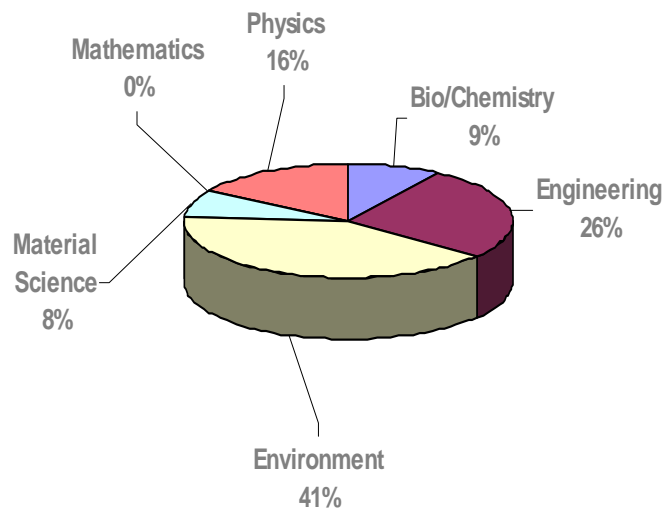


PVP NEC SX-5

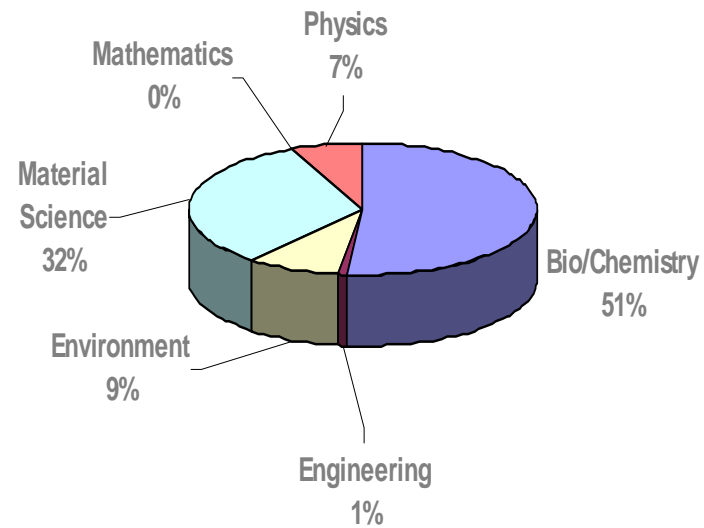


Usage by field, 2003

PVP NEC SX-5



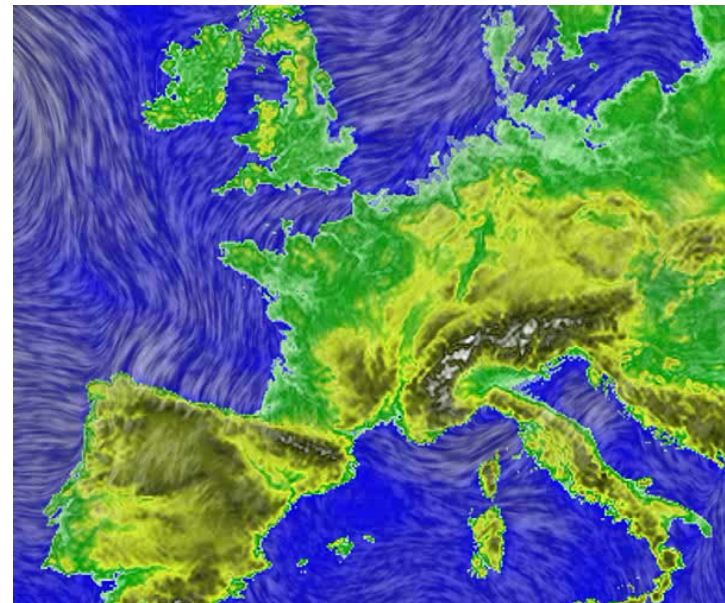
MPP IBM Power4



Meteo and clima modelling framework



Each day, MeteoSwiss produces its numerical weather predictions, including daily operational short range weather forecasts and research activities are run at CSCS.



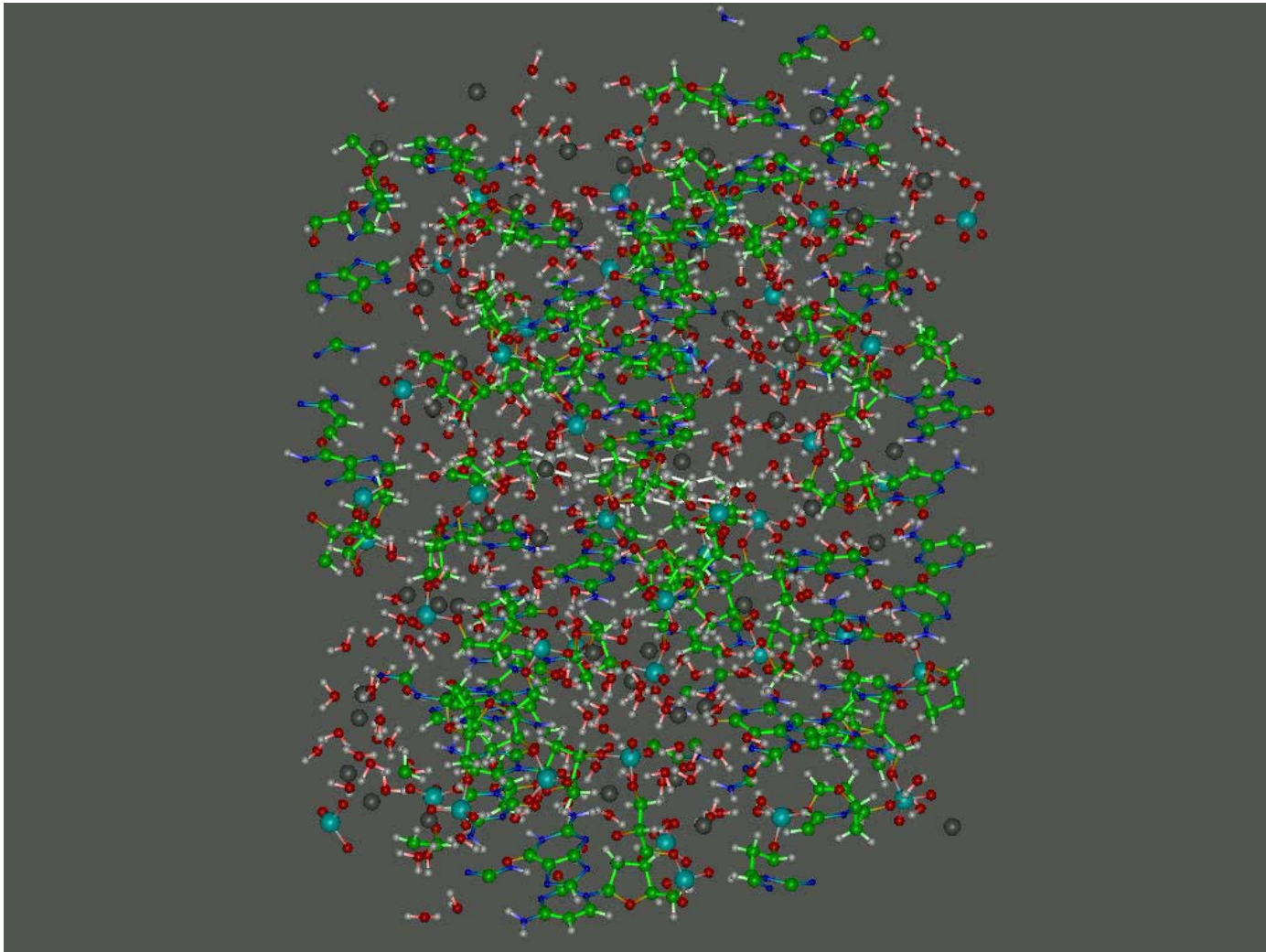
EUROGRID



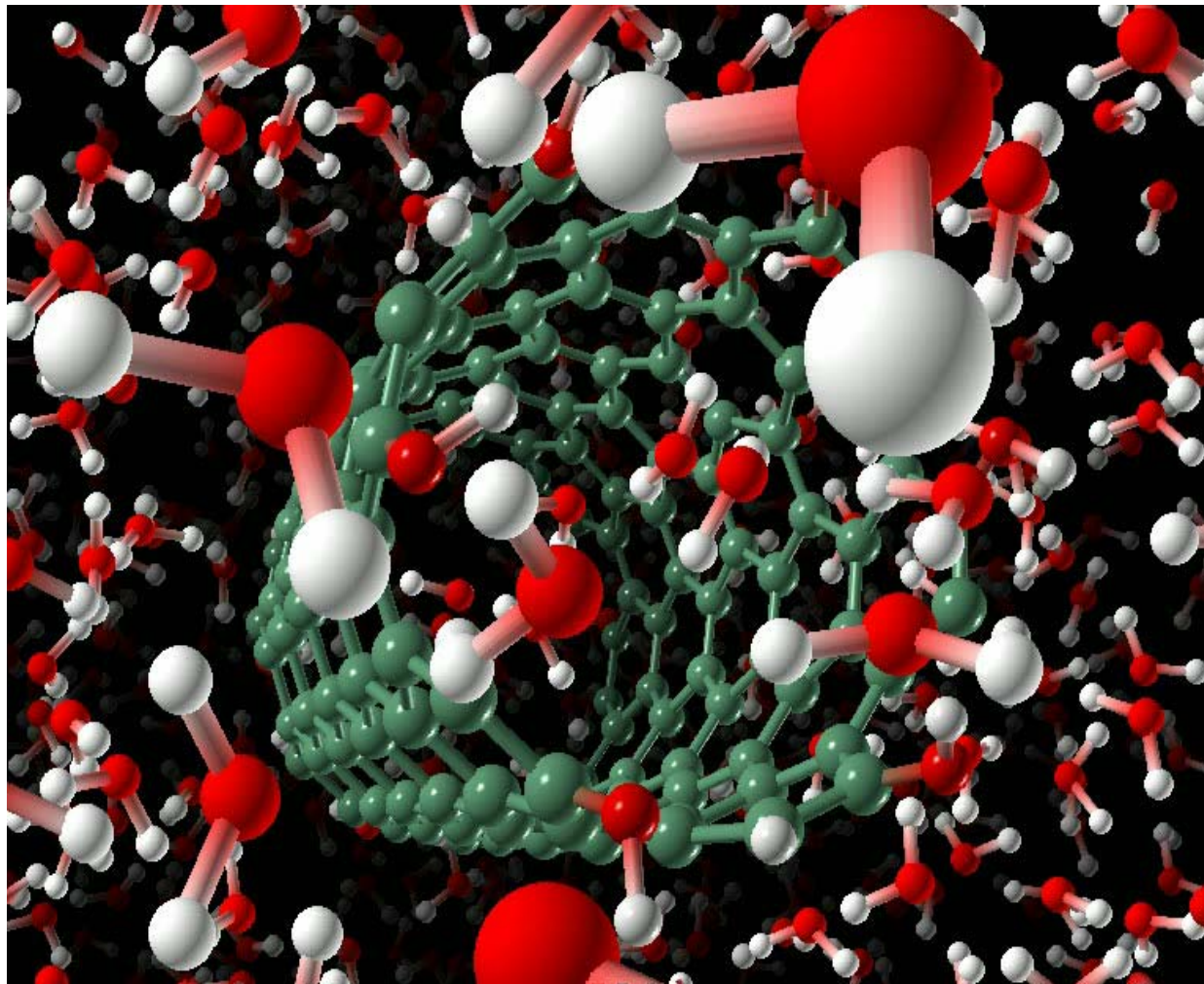
Major EU financed R&D and SW engineering undertaking to develop important Grid software components (Unicore) and to integrate, operate and support them into EUROGRID software infrastructure operated and supported by the leading High Performance Computing centres from different European countries. After project end the EUROGRID software will be available as a supported product.



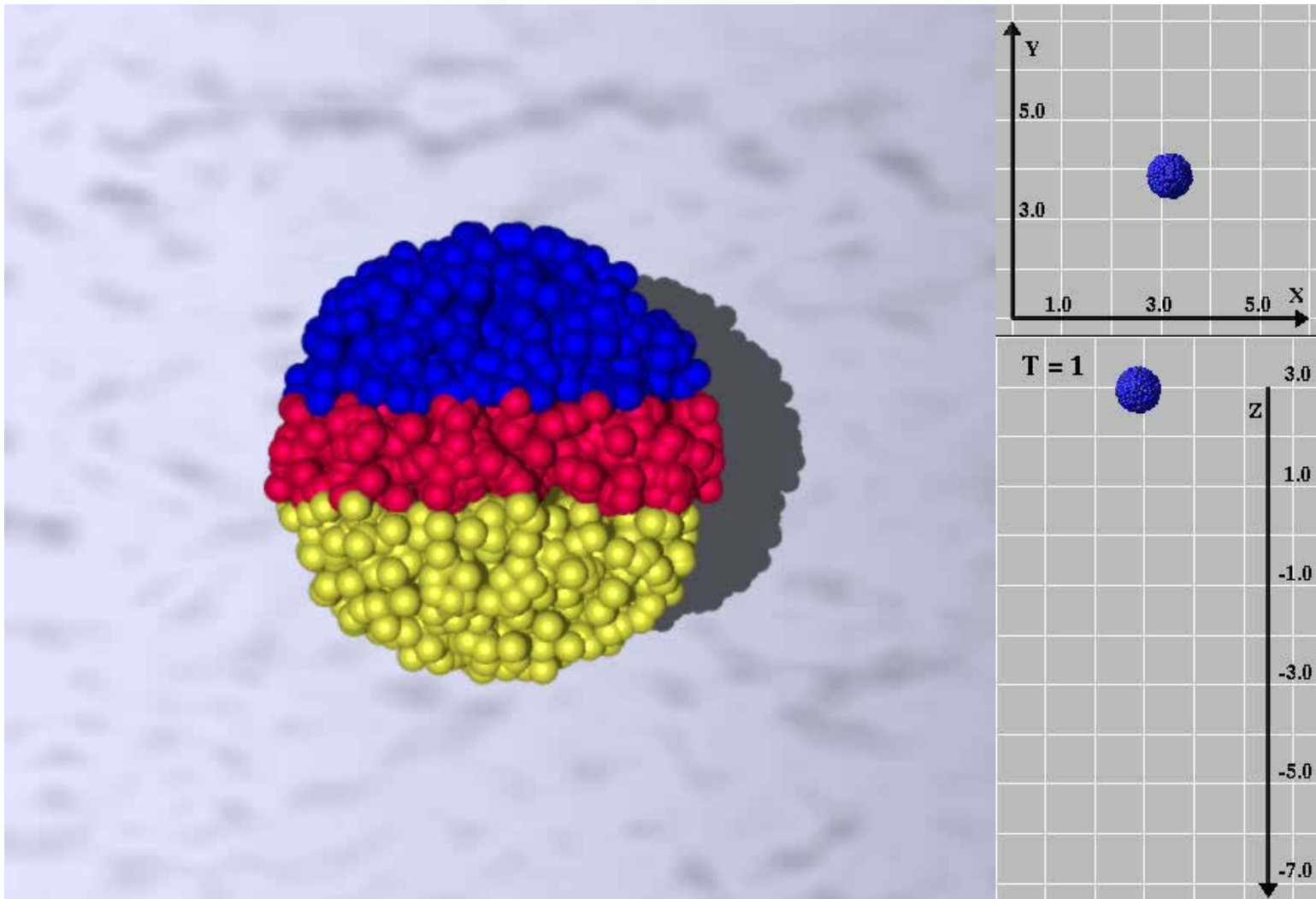
Example of scientific visualization



Graphics and visualization



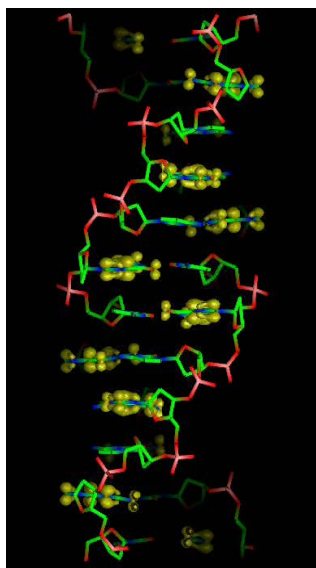
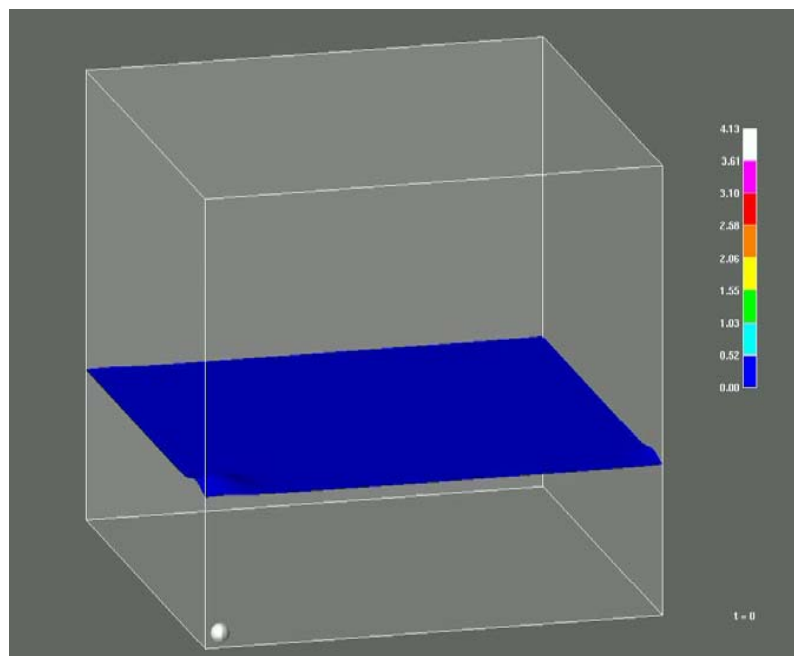
Example of scientific visualization



Examples of graphics applications at CSCS



Dianaline



High Orbital
Molecular Occupancy

Driving forces

Capability

Ability to serve the projects of the highest scientific quality requiring intensive resources

Sustainability

Clear line of funding, and economy of scale

Scalability

Streamlining a set of services from the desktop to highest end servers

Flexibility

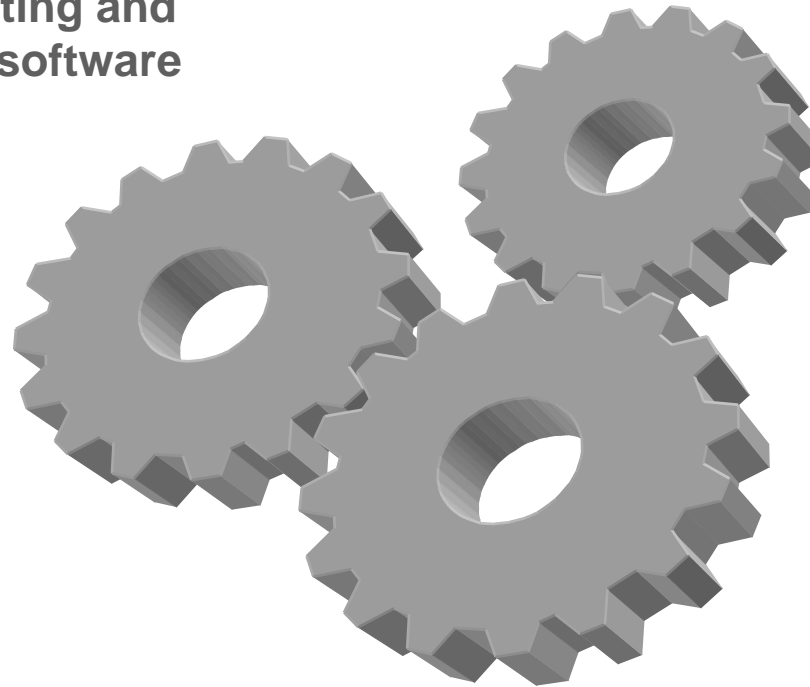
Capacity to evolve according to the pace of technology

A strategy based on 3 goals



Complementary HPCN computer architectures

**High Performance Computing and
Networking methods and software**



Positioning the centre both nationally and internationally by partnering with key players

CSCS

mission statement



CSCS is the Swiss National Supercomputing Centre, providing, developing and promoting technical and scientific services for the Swiss research community on the fields of high-performance and high-throughput computing. It is a centre of competence that pioneers new information technologies; collaborates with domestic and foreign researchers, and carries out its own research and development in computational sciences and scientific computing.

Scientific portfolio



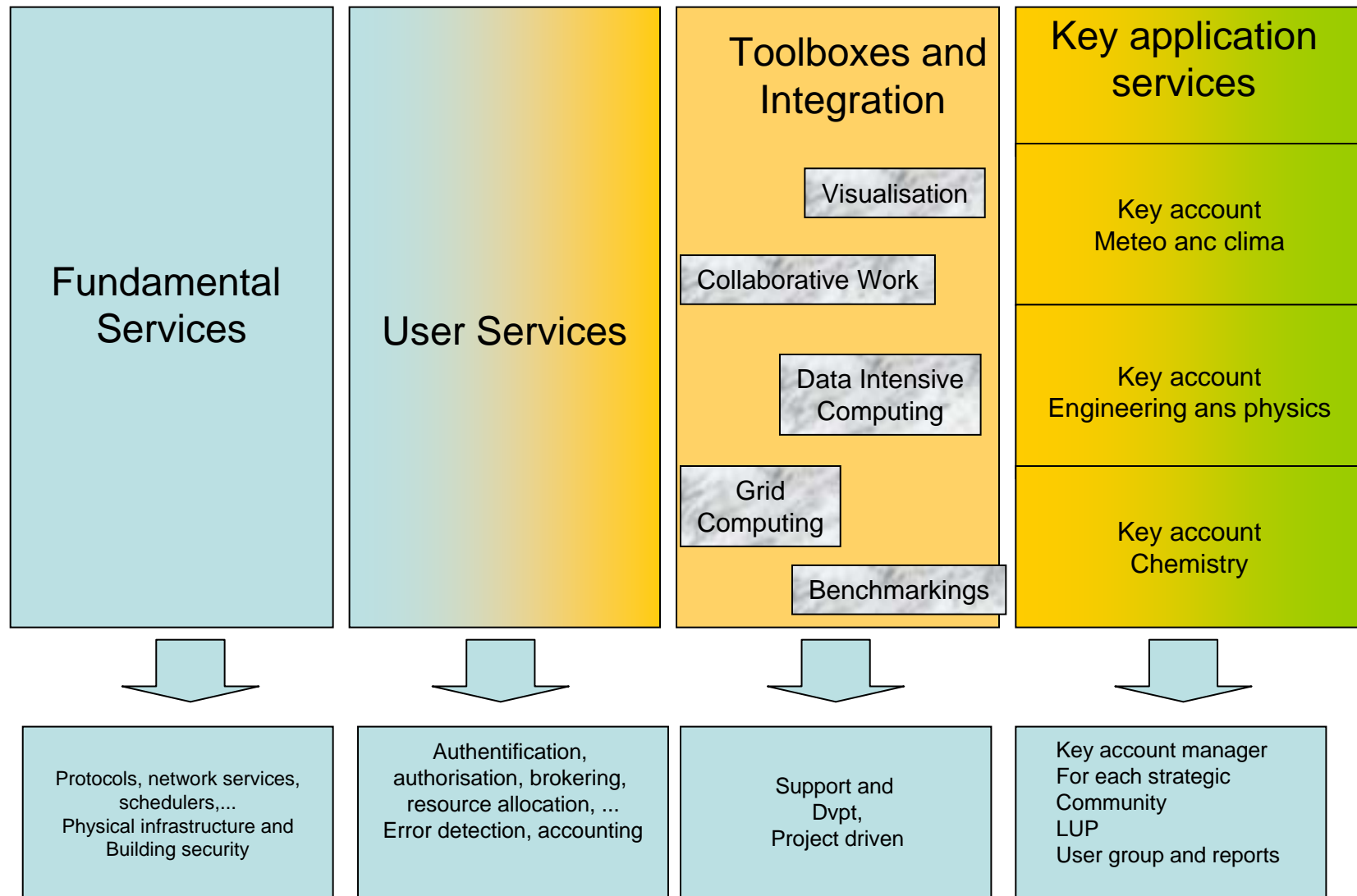
	Application Field	CSCS Contribution
reinforce	Engineering, Computational Chemistry, Physics	software framework and expertise in application optimisation
	Meteorology and Climate Modelling	NCCR Klima, contract work for Meteo Swiss, Prism collaboration
	Material and Molecular Sciences	Expertise in application optimisation, plus hosting CPMD-software & support to the Parrinello group
gradually develop	Grid Computing	towards e-science, with ETHZ EPFL, USI, Unis and key players
	Earth Science & Global Modelling	software framework for supporting community
	Life Sciences	development of visualization/imaging techniques
	Information Science	hosting of data and development of filtering and retrieval methods

New thrust areas of scientific computing



	Discipline	Description
reinforce	Visualization	Imaging, virtual reality,
	Parallelism, Numerics & Benchmarking	Optimisation, selection and development of solvers and other mathematical methods
	Data Intensive Computing	Data mining, knowledge management, information management and retrieval, machine learning, geographical information systems
gradually develop	Distributed Computing	Grid computing, networking, communications, agents, distributed services
	Education & Training	Organisation of courses and summer schools, visiting scientists
	Modelling Framework Support	Hosting and fostering a software framework that is being used as the community model of a research discipline

Services organization

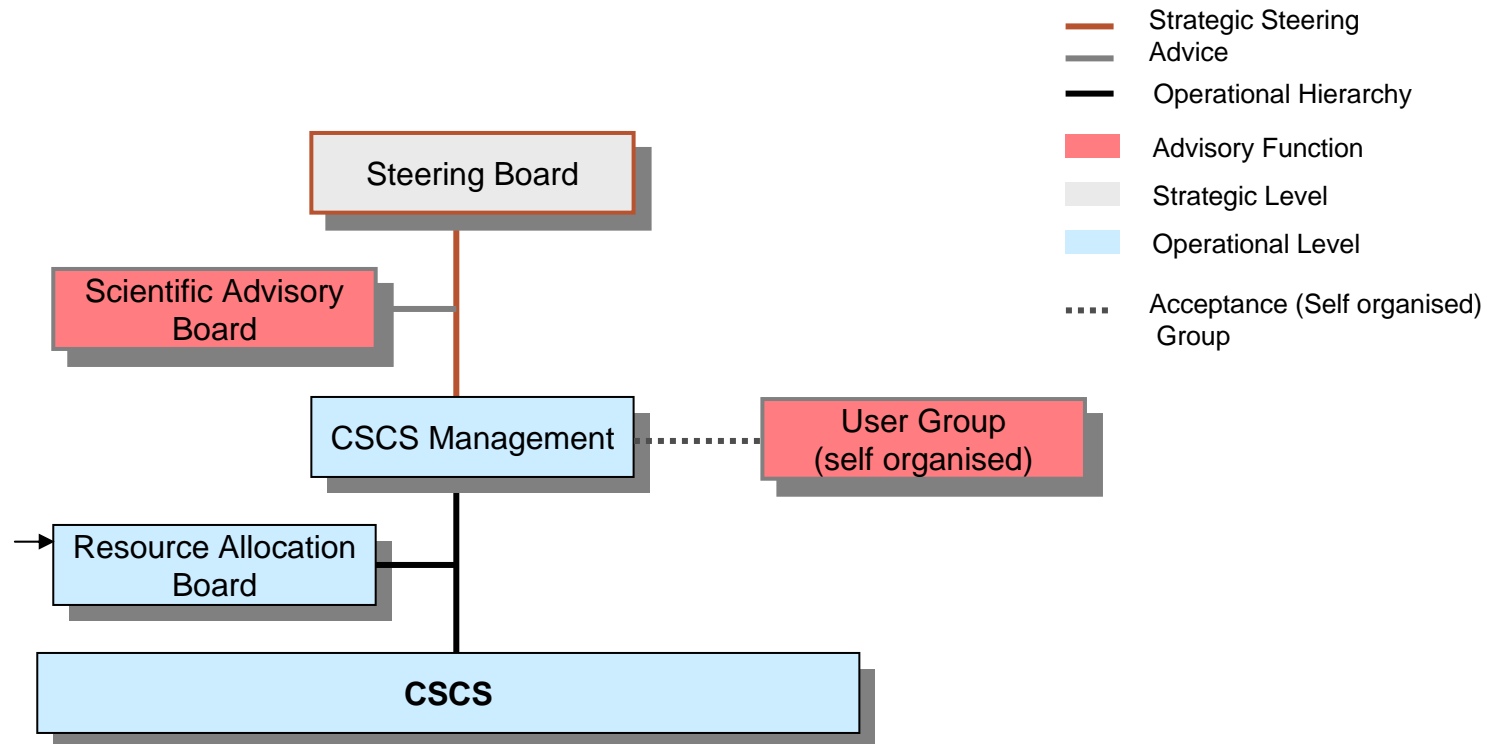


CSCS Autonomy



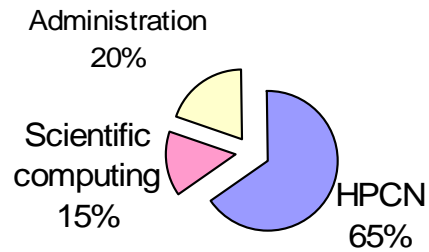
- Unit of ETHZ managed under performance mandate and global budget
- Similar to FLAG institutions

Governance

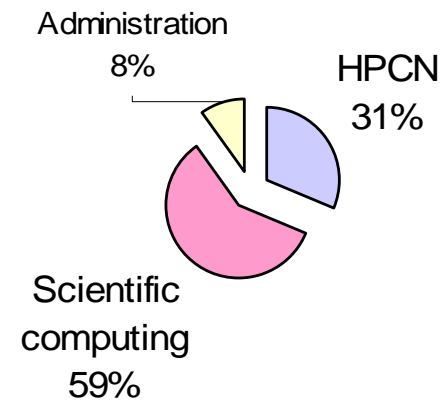


Expertise development at CSCS

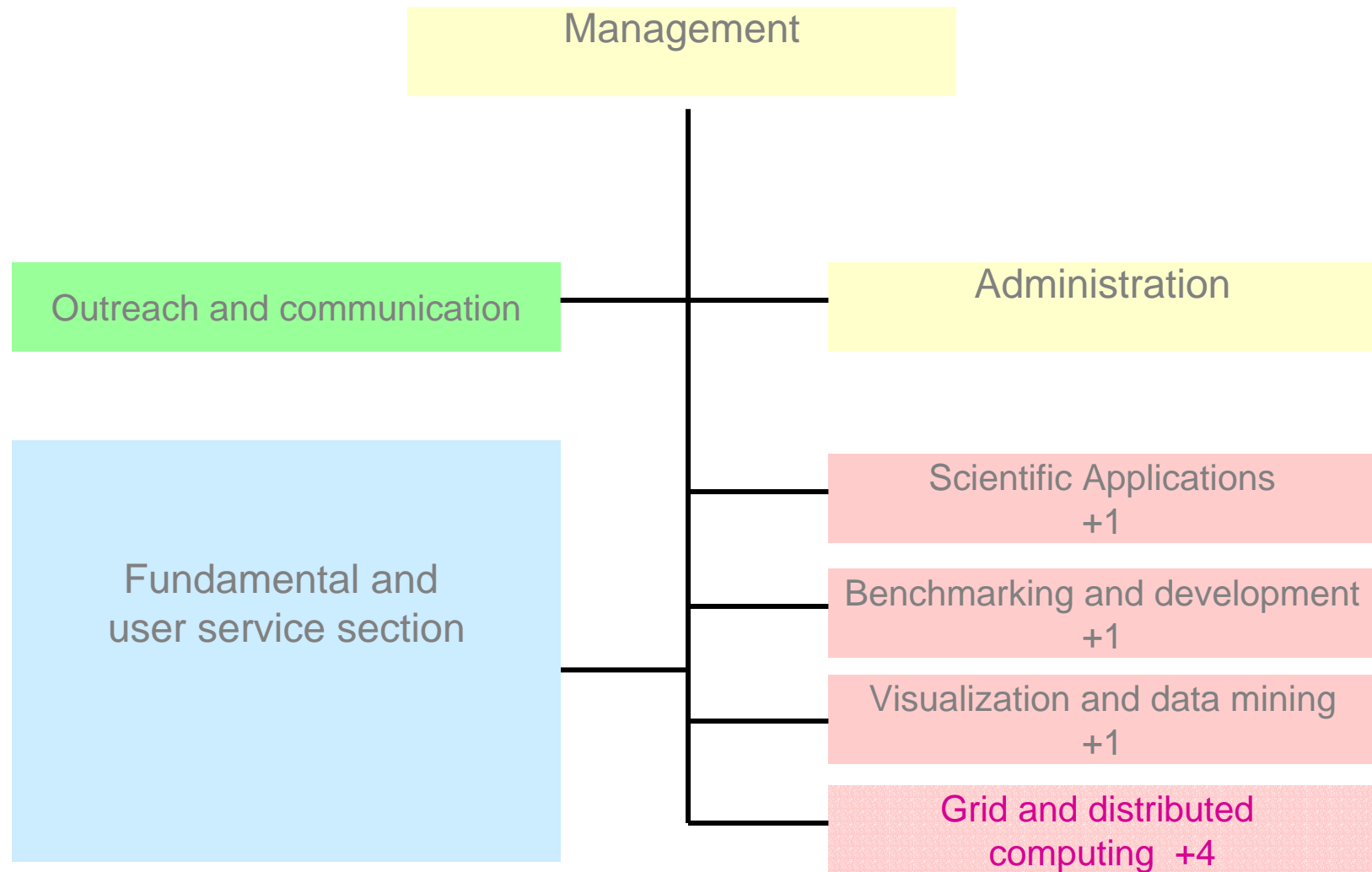
CSCS Today



CSCS >2003



CSCS organigramm



A short history of CSCS



1991- 1999

National service with a wide customer base

Since 2000

National service with strong leadership in
computational science

2004-2010

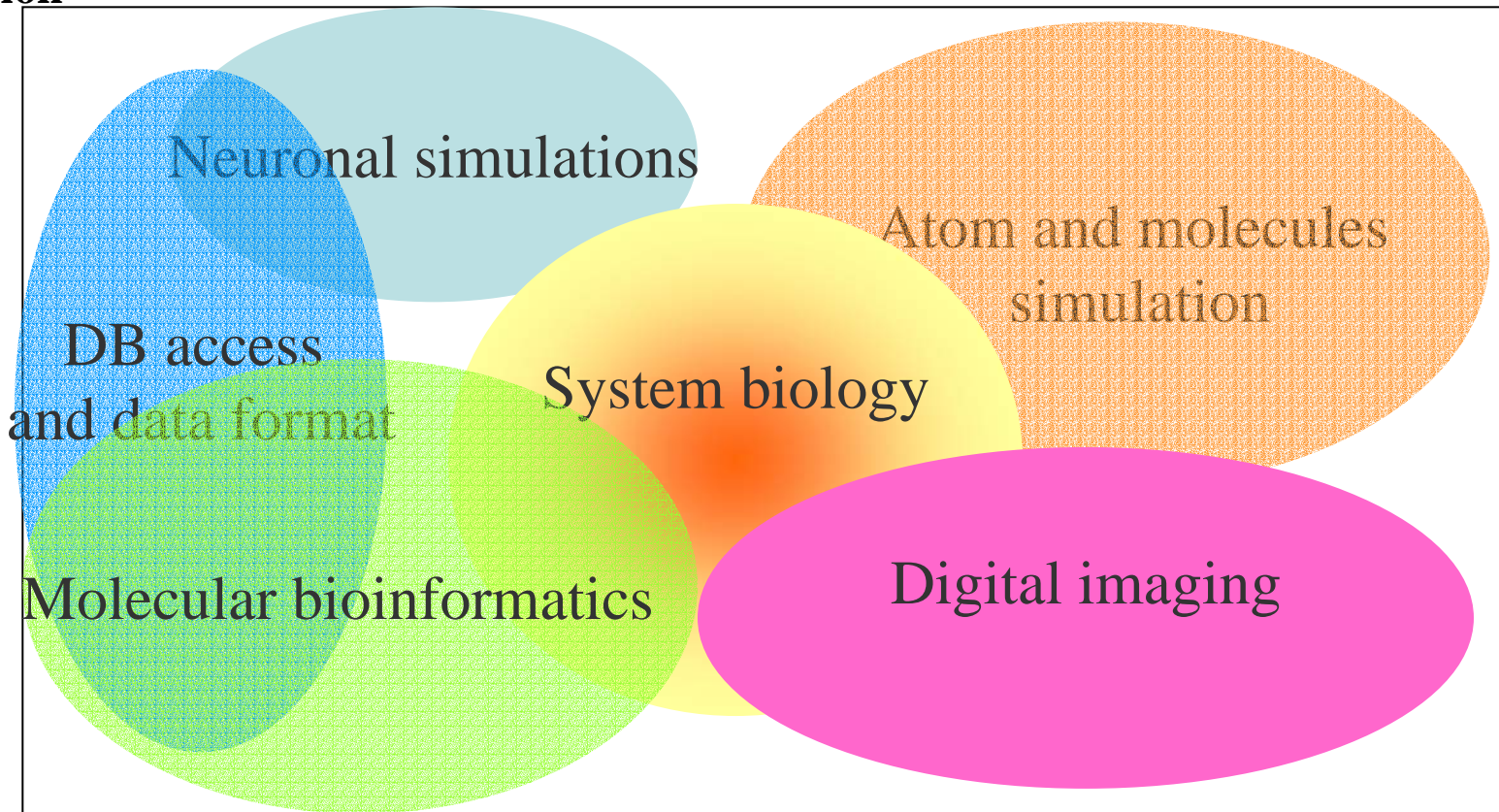
National service with strong leadership for IT and
scientific computing support and development

Grid and distributed computing: first actions

- [LCG Grid](#): 7 institutes, CERN and CSCS
- [Swiss BioGrid](#): 5 institutes including one Pharma
- *Swiss Computational Grid: ETHZ, EPFL, PSI, CSCS, UniGe, Unibe, Switch,...*

Integrative map for life sciences

Modelisation



Analyse

Data intensive

Compute intensive

Grids as collaborative platforms



Grille de Calcul

soutien à l'interdisciplinarité et à la constitution du savoir dans les sciences du vivant



Marie-Christine.Sawley@epfl.ch, VPR-DAV-EPFL

L'information, composante du portefeuille de valeurs

La biologie et les sciences du vivant sont en pleine mutation, évoluant de sciences de l'observation vers une discipline prédictive et systématique. L'énorme quantité d'informations générée nécessite une infrastructure informatique de plus en plus performante pour extraire l'information scientifique rapidement. Cette constatation touche par exemple des domaines aussi divers que l'imagerie médicale, le séquençage et l'annotation du génome. Il est essentiel de pouvoir acquérir, traiter, stocker, récupérer, analyser et diffuser cette information de la meilleure façon possible.

Les centres de calcul actuels ne peuvent faire face seuls à cette croissance exponentielle en capacité de traitement ni à la demande d'absorption des pointes ou de flexibilité d'évolution des infrastructures. Comme toute période de mutation, celle-ci engendre risque et occasions de développement.

Le bassin de compétences, garant de la masse critique

Dans un rayon de moins de 100 km, notre région constitue un *réseau* offrant un large éventail d'avantages complémentaires pour le développement de l'interface entre l'informatique et les sciences du vivant: la présence de

- trois centres universitaires et de deux hôpitaux cantonaux. Au total, ce sont plus de 11'000 étudiants qui sont inscrits en sciences de l'ingénieur à l'EPFL et dans les facultés de sciences et de médecine des universités de Genève et Lausanne. Une proportion importante des chercheurs académiques à l'EPFL exerce dans des domaines ayant des ramifications interdisciplinaires vers les sciences du vivant.
- deux centres de recherche de très grande renommée (ISREC et LUDWIG) et l'Institut Suisse de Bioinformatique (SIB). Ils ont bâti une compétence importante, gèrent et enrichissent d'importantes bases de données (protéomique, acides nucléiques,...) qu'ils mettent à disposition de la communauté scientifique internationale, et totalisent plus de 240 chercheurs.
- un environnement entrepreneurial en biotechnologie: Serono, Cresogen, Moden, Debsopharm, Debio Recherche, Baxter et Diamid.
- un des leaders mondiaux du biomédical (Medtronic).
- plusieurs incubateurs et pépinières d'entreprises.
- une forte présence de Nestlé, Firmenich, Givaudan, Dupont.

- l'implantation de plusieurs grandes sociétés informatiques
- des ressources très performantes et des compétences de pointe pour les télécommunications.
- la proximité du Cern, *«sovereign site Web sans horde»*.

UNE POLITIQUE DE PARTENARIATS INDISPENSABLE

La constitution du savoir scientifique dans ces nouvelles disciplines et ses retombées importantes en matière de formation, d'innovations et de transfert technologique et à terme, de création d'emplois, nécessite un effort très important que seule une politique de partenariats choisis peut soutenir. La région lémanique offre un cadre particulièrement bien adapté à cet essor.

D'une part, les entreprises multiplient les sources pour identifier et saisir de la chaîne d'innovations des opportunités de développement dans des niches spécialisées: intramuros, auprès des *startups* ou, à un stade plus précoce encore, dans les laboratoires de recherche fondamentale. Les relations avec le monde de la recherche académique sont en pleine évolution.

Les *startups* technologiques, qui sont parfois issues de cette interface, ont d'ailleurs un rôle particulier et des besoins propres en terme d'accès à l'infrastructure et au réseau de communications.

Enfin, un mouvement amorce il y a quelques années dans le secteur universitaire a franchi l'été dernier une étape importante. L'Université de Genève, l'EPFL et l'Université de Lausanne ont affirmé clairement leur volonté d'aller au-delà d'une simple coordination en mettant en place les meilleures conditions pour des collaborations inter-institutionnelles, pluridisciplinaires, à forte valeur ajoutée pour la jeunesse de notre pays, comme l'atteste la signature de la convention du programme *Science, Vie, Société*.

UN PRE

Parmi les sciences du tic: discipline de l'analyse de l'information biologique, en majorité sous la forme de séquences génétiques et de structures de protéines, les premiers travaux sont antérieurs à la révolution géomique même si le terme *bioinformatique* n'est apparu qu'il y a une dizaine d'années.

January 2002

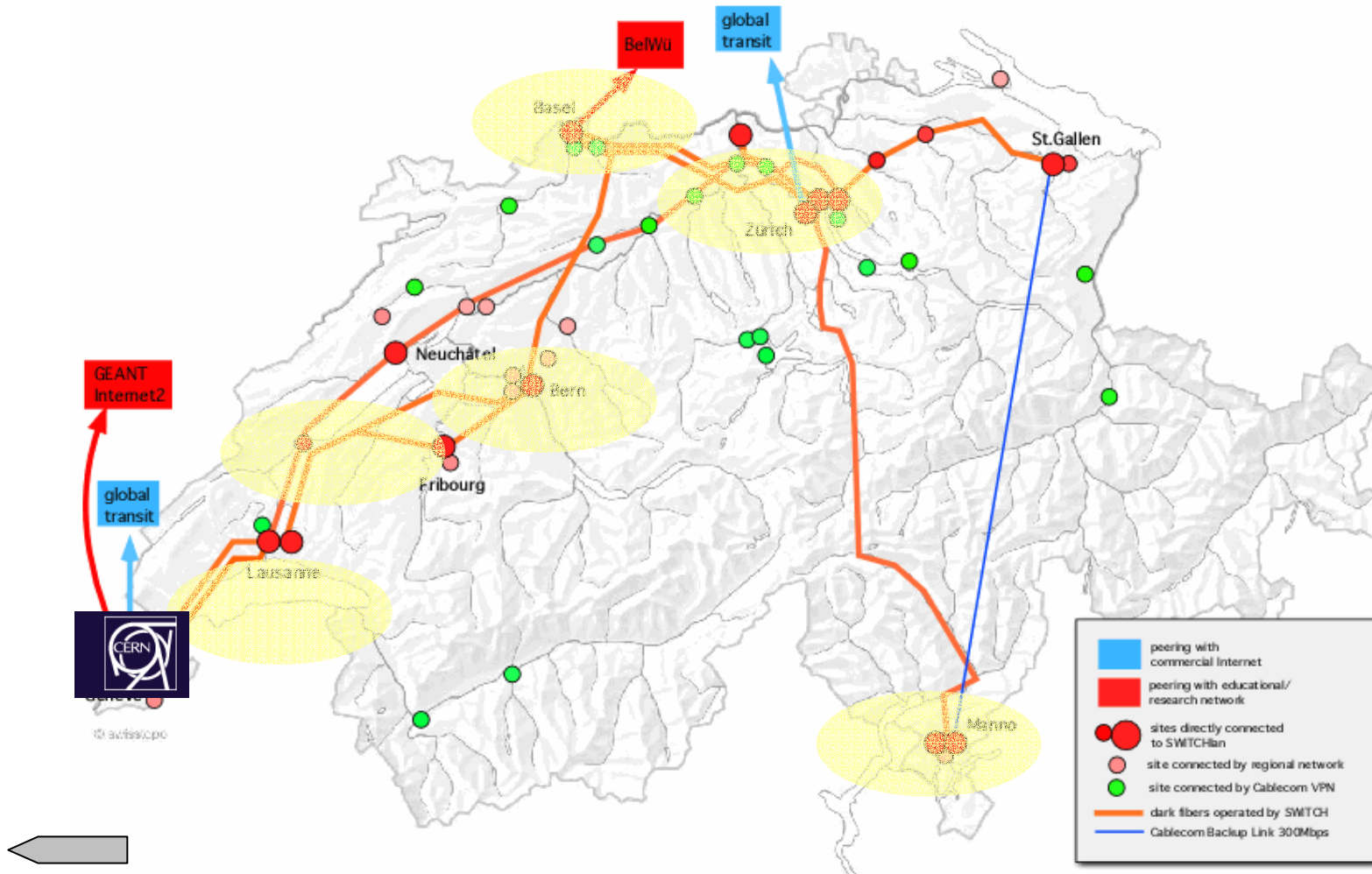
"...we see a major growth of volume of scientific data

1. results of simulation
2. experimental data
3. documents: texts, images, videos
4. tutorials and teaching material
5. software and modules

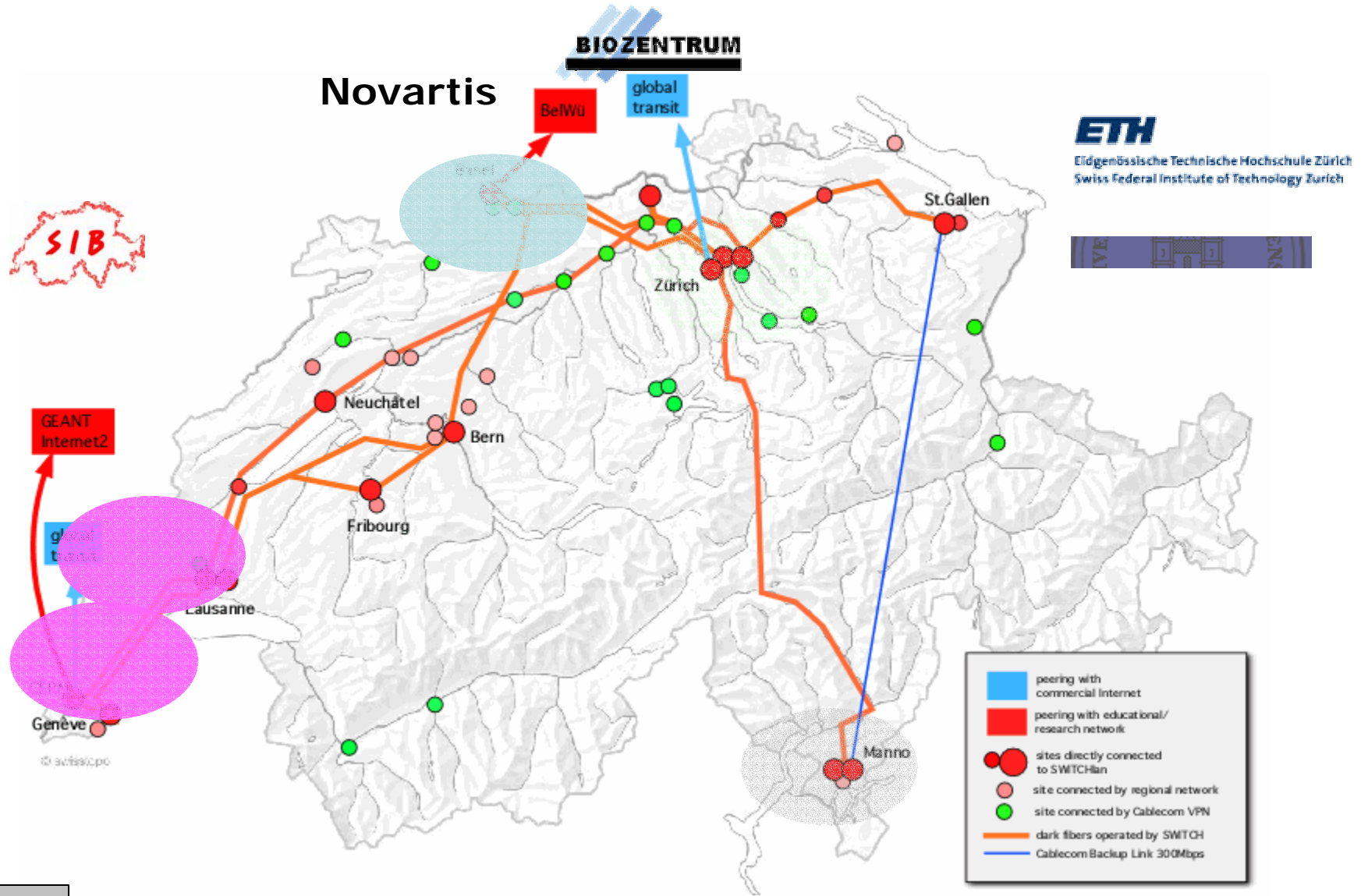
The need to automate the process of going from raw information originating from different groups to **structured knowledge accessible to multi-disciplinary, multi-lingual, multi-institutional communities, is emerging**".

March 2002

LCG in CH



Swiss Bio Grid



Goals of the Swiss Bio Grid



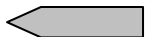
- *Enables and stimulates interdisciplinary research, therefore enhances the capacity of scientific production;*
- *allows the tackling of computational problems of an order of magnitude higher than those presently addressed;*
- *stimulates the development of IT literacy (know-how and knowledge) that benefit other areas of the academic or industrial sector: “learning together for increasing knowledge” ;*
- *optimize resource utilization;*

Swiss BioGrid

role of the CSCS



- authentication and security over the SBG
- management tools for accessing the resources and clearing
- compute and data grid, by making part its resource available and hosting part of the IT systems
- knowledge management, by exploring, testing and validating the tools, integrated into a portal dedicated to the Bio sciences



> 2004

additional servers and networking capabilities

- Goals
 - Enforcement of policy-based management for resource allocation
 - Knowledge building and dissemination: e-science portal
 - Funding increase from non-governmental sources
- Action items
 - Prototyping and implementation of specific grid activities such as:
 - Computing on demand and computational steering, remote visualization, collaborative tools, remote instrumentation, brokering agents, etc.
 - Supporting and developing e-science testbeds

Goals of the Swiss HPCN Grid leading house



- To build a network of Scientific centres and “Regional Grids”
 - Establish a Grid Network team, also with Switch
- Continue to initiate Thematic Projects
- Support for e-Science Pilot Projects
 - Scholarships for a limited number for PhD students or post docs
- Push further research boundaries:
 - the imagination of the scientists is the ceiling!
- Promote connectivity between disciplines and users
 - Knowledge management
- Reinforce international collaborations.

CSCS Role for Middleware Infrastructure

- Repository for open source, standard compliant, middleware stack
 - Knowledge management, training and outreach
- Software Engineering as well as R&D
 - to produce robust, professionally documented, re-usable software
 - to embrace emerging Grid Service standards
 - collaboration with Universities of Applied Sciences and incubators