Activity Report 2017

Section Application Domains

ASAP Project-Team (section vide)

COAST Project-Team (section vide)

<h1 style="text-align:center; color:red;">CTRL-A Project-Team</h1>

4. Application Domains

4.1. Self-adaptive and reconfigurable computing systems in HPC and the IoT

We are attacking the problem of designing well-regulated and efficient self-adaptive computing systems by the development of novel strategies for systems management.

The kind of systems we typically target involve relatively coarse grained computation tasks (e.g. image processing or HPC tasks, components or services), assembled in workflows, application dependency graphs, or composites. At that level, there can be parallel and conditional branches, as well as choices that can be made between alternative branches, corresponding to different ways to perform that part of the application. Such tasks can be achieved following a choice of implementations or versions, such as in service oriented approaches. Each implementation has its own characteristics and requirements, e.g., w.r.t. resources consumed and QoS offered. The systems execution infrastructures present heterogeneity, with different computing processors, a variety of peripheral devices (e.g., I/O, video port, accelerators), and different means of communications. This hardware or middleware level also presents adaptation potential e.g., in varying quantities of resources or sleep and stand-by modes.

The kinds of control problems encountered in these adaptive systems concern the navigation in the configurations space defined by choice points at the levels of applications, tasks, and architecture. Upon events or conditions triggering reconfiguration and adaptation, the controller has to choose a next configuration where, on the one hand, all consistency constraints are satisfied w.r.t. dependencies and resources requirements. On the other hand, it has to apply a policy or strategy deciding between correct configurations which one to chose e.g. by optimizing one or more criteria, or by insuring reachability of some later configuration (goal or fallback). This targeted class of computing systems we consider is mid-sized, in the sense that the combinatorial complexity is large enough for manual solving to be impossible, while remaining within the range where supervisory control techniques are tractable. The pace of control is more sporadic, and slower than the instruction-level computation performance within the coarse-grained tasks.

The objectives of CTRL-A will be achieved and evaluated in both of our main application domains, thereby exhibiting their similarities from the point of view of reconfiguration control. A first application domain is High Performance Computing. In this area, we currently focus especially on the management of Dynamic Partial Reconfiguration in FPGA based hardware, at the level of middleware. Here the particular application we consider is, as in our ANR project HPeC starting end of 2015, video image flow processing for smart cameras implemented on DPR FPGASs themselves embedded in drones.

A second general application domain to confront our models is the Internet of Things (IoT), more specifically self-adaptive middleware platforms for Smart Environments, or Industry 4.0 related topics, like SCADA. We focus on providing coordination components and controllers of software components and services, or rule-based middleware platforms. The adaptation problems concern both the functional aspects of the applications in a smart building, and the middleware support deployment and reconfiguration issues. We are considering perspectives concerning self-protection and security.

MIMOVE Team

4. Application Domains

4.1. Mobile urban systems for smarter cities

With the massive scale adoption of mobile devices and further expected significant growth in relation with the Internet of Things, mobile computing is impacting most – if not all – the ICT application domains. However, given the importance of conducting empirical studies to assess and nurture our research, we focus on one application area that is the one of "*smart cities*". The smart city vision anticipates that the whole urban space, including buildings, power lines, gas lines, roadways, transport networks, and cell phones, can all be wired together and monitored. Detailed information about the functioning of the city then becomes available to both city dwellers and businesses, thus enabling better understanding and consequently management of the city's infrastructure and resources. This raises the prospect that cities will become more sustainable environments, ultimately enhancing the citizens' well being. There is the further promise of enabling radically new ways of living in, regulating, operating and managing cities, through the increasing active involvement of citizens by ways of crowd-sourcing/sensing and social networking.

Still, the vision of what smart cities should be about is evolving at a fast pace in close concert with the latest technology trends. It is notably worth highlighting how mobile and social network use has reignited citizen engagement, thereby opening new perspectives for smart cities beyond data analytics that have been initially one of the core foci for smart cities technologies. Similarly, open data programs foster the engagement of citizens in the city operation and overall contribute to make our cities more sustainable. The unprecedented democratization of urban data fueled by open data channels, social networks and crowd sourcing enables not only the monitoring of the activities of the city but also the assessment of their nuisances based on their impact on the citizens, thereby prompting social and political actions. However, the comprehensive integration of urban data sources for the sake of sustainability remains largely unexplored. This is an application domain that we intend to focus on, further leveraging our research on emergent mobile distributed systems, large-scale mobile sensing & actuation, and mobile social crowd-sensing.

In a first step, we concentrate on the following specialized applications, which we investigate in close collaboration with other researchers, in particular as part of the dedicated Inria Project Lab *CityLab@Inria*:

– **Democratization of urban data for healthy cities.** The objective here is to integrate the various urban data sources, especially by way of crowd-Xing, to better understand city nuisances from raw pollution sensing (e.g., sensing noise) to the sensing of its impact on citizens (e.g., how people react to urban noise and how this affects their health).

– **Socially-aware urban mobility.** Mobility within mega-cities is known as one of the major challenges to face urgently due to the fact that today's mobility patterns do not scale and to the negative effect on the environment and health. It is our belief that mobile social and physical sensing may significantly help in promoting the use of public transport, which we have started to investigate through empirical study based on the development and release of dedicated apps.

– **Social applications.** Mobile applications are being considered by sociologists as a major vehicle to actively involve citizens and thereby prompt them to become activists. This is especially studied with the Social Apps Lab at UC Berkeley. Our objective is to study such a vehicle from the ICT perspective and in particular elicit relevant middleware solutions to ease the development and development of such "*civic apps*".

Acknowledging the need for collaborative research in the application domain of smart cities, MiMove is heavily involved and actually leading CityLab@Inria [0]. The Inria Project Lab CityLab is focused on the study of ICT solutions promoting social sustainability in smart cities, and involves the following Inria project-teams in addition to MiMove: CLIME, DICE, FUN, MYRIADS, SMIS, URBANET and WILLOW. CityLab further involves strong collaboration with California universities affiliated with CITRIS (Center for Information Technology Research in the Interest of Society) and especially UC Berkeley, in relation with the *Inria@SiliconValley* program. We note that Valérie Issarny acts as scientific manager of Inria@SiliconValley and is currently visiting scholar at CITRIS at UC Berkeley. In this context, MiMove researchers are working closely with colleagues of UC Berkeley, including researchers from various disciplines interested in smart cities (most notably sociologists).

4.2. Home Network Diagnosis

With the availability of cheap broadband connectivity, Internet access from the home has become a ubiquity. Modern households host a multitude of networked devices, ranging from personal devices such as laptops and smartphones to printers and media centers. These devices connect among themselves and to the Internet via a local-area network —a *home network*– that has become an important part of the "Interne experience". In fact, ample anecdotal evidence suggests that the home network can cause a wide array of connectivity impediments, but their nature, prevalence, and significance remain largely unstudied.

Our long-term goal is to assist users with concrete indicators of the causes of potential problems and—ideally—ways to fix them. We intend to develop a set of easy-to-use home network diagnosis tools that can reliably identify performance and functionality shortcomings rooted in the home. The development of home network diagnosis tools brings a number of challenges. First, home networks are heterogeneous. The set of devices, configurations, and applications in home networks vary significantly from one home to another. We must develop sophisticated techniques that can learn and adapt to any home network as well as to the level of expertise of the user. Second, there are numerous ways in which applications can fail or experience poor performance in home networks. Often there are a number of explanations for a given symptom. We must devise techniques that can identify the most likely cause(s) for a given problem from a set of possible causes. Third, even if we can identify the cause of the problem, we must then be able to identify a solution. It is important that the output of the diagnosis tools we build is "actionable". Users should understand the output and know what to do.

We are conceiving methods for two application scenarios: (i) when the end user in the home deploys our diagnostic tools either on the home gateway (the gateway often combines a DSL/cable modem and an access point; it connects the home network to the ISP) or on devices connected to the home network and (ii) when ISPs collect measurements from homes of subscribers and then correlate these measurements to help identify problems.

Assisting end users. We are developing algorithms to determine whether network performance problems lie inside or outside the home network. Given that the home gateway connects the home with the rest of the Internet, we are designing an algorithm (called *HoA*) that analyzes traffic that traverses the gateway to distinguish access link and home network bottlenecks. A measurement vantage point on the gateway is key for determining if the performance bottleneck lies within the home network or the access ISP, but we also need to deploy diagnosis tools in end-devices. First, some users may not want (or not know how) to deploy a new home gateway in their homes. Second, some problems will be hard to diagnose with only the vantage point of the gateway (for example, when a device cannot send traffic or when the wireless is poor in certain locations of a home). We can obtain more complete visibility by leveraging *multiple* measurement nodes around the home, potentially including the home gateway, all participating jointly in the measurement task. We have an ongoing project to realize a home network analyzer as a web-based measurement application built on top of our team's recently developed browser-based measurement platform, *Fathom*. To integrate the home gateway in the analyzer, we plan to engage the BISmark Project. BISmark already provides a web server as well as

---

extensive configurability, allowing us to experiment freely with both passive as well as active measurements. We must develop a home network analyzer that can first discover the set of devices connected to the home network that can collaborate on the diagnosis task. We will then develop tomography algorithms to infer where performance problems lie given measurements taken from the set of available vantage points.

Assisting Internet Service Providers (ISPs). Our discussions with several large access ISPs reveal that service calls are costly, ranging from $9–25 per call, and as many as 75% of service calls from customers are usually caused by problems that have nothing to do with the ISP. Therefore, ISPs are eager to deploy techniques to assist in home network diagnosis. In many countries ISPs control the home gateway and set-top-boxes in the home. We plan to develop more efficient mechanisms for home users to report trouble to their home ISP and consequently reduce the cost of service calls. This project is in collaboration with Technicolor and Portugal Telecom. Technicolor is a large manufacturer of home gateways and set-top-boxes. Portugal Telecom is the largest broadband access provider in Portugal. Technicolor already collects data from 200 homes in Portugal. We are working with the data collected in this deployment together with controlled experiments to develop methods to diagnose problems in the home wireless.

4.3. Quality of Experience

An increasing number of residential users consume online services (e.g., VoD, Web browsing, or Skype) in their everyday activities (e.g., for education or entertainment purposes), using a variety of devices (e.g., tablets, smartphones, laptops). A high Quality of Service (QoS) is essential for sustaining the revenue of service providers, carriers, and device manufactures. Yet, the perceived Quality of Experience (QoE) of users is far from perfect e.g., videos that get stalled or that take a long time to load. Dissatisfied users may change Internet Service Providers (ISPs) or the online services. Hence, the incentives for measuring and improving QoE in home networks are high while mapping network and application QoS to QoE is a challenging problem. In this work we have focused in measuring several network Quality-of-Service (QoS) metrics, such as latency and bandwidth, both in residential Wi-Fi as well as broadband networks, homes are using for connecting to the Internet.

The WiFi Context. Residential Wi-Fi performance, however, is highly variable. Competing Wi-Fi networks can cause contention and interference while poor channel conditions between the station and the access point (AP) can cause frame losses and low bandwidth. In some cases, the home Wi-Fi network can bottleneck Internet access. While problems in the Wi-Fi network may affect several network QoS metrics, users will typically only notice a problem when poor Wi-Fi affects the QoE of Internet applications. For example, a Wi-Fi network with low bandwidth may go unnoticed unless the time to load Web pages increases significantly. A user observing degraded QoE due to Wi-Fi problems may mistakenly assume there is a problem with the Internet Service Provider (ISP) network. Our discussions with residential ISPs confirm that often customers call to complain about problems in the home Wi-Fi and not the ISP network.

Prior work has focused on QoS metrics for some applications (e.g., on-line video, Web browsing, or Skype) with no attempt to identify when Wi-Fi quality affects QoE. We are particularly interested in assisting ISPs to predict when home Wi-Fi quality degrades QoE. ISPs can use this system to detect customers experiencing poor QoE to proactively trigger Wi-Fi troubleshooting. ISPs often control the home AP, so we leverage Wi-Fi metrics that are available on commercial APs. Detecting when Wi-Fi quality degrades QoE using these metrics is challenging. First, we have no information about the applications customers are running at any given time. ISPs avoid capturing per-packet traffic traces from customers, because of privacy considerations and the overload of per-packet capture. Thus, we must estimate the effect of Wi-Fi quality on QoE of popular applications, which most customers are likely to run. In this context, we study Web as a proof of concept, as a large fraction of home traffic corresponds to Web. Second, application QoE may be degraded by factors other than the Wi-Fi quality (e.g., poor Internet performance or an overloaded server). Although a general system to explain any QoE degradation would be extremely helpful, our monitoring at the AP prevents us from having the end-to-end view necessary for such general task. Instead, we focus on identifying when Wi-Fi quality degrades QoE. Finally, Wi-Fi metrics available in APs are coarse aggregates such as the average PHY rate or the fraction of busy times. It is open how to effectively map these coarse metrics into QoE.

Predicting QoE. Clearly, different actors in the online service chain (e.g., video streaming services, ISPs) have different incentives and means to measure and affect the user QoE. Uncovering statistically equivalent subsets of QoS metrics across and within levels provides actionable knowledge for building QoE predictors. To achieve this goal, we leverage recent advances on feature selection algorithms to exploit available experimental evidence of the joint probability distributions of QoE/QoS metrics. This type of statistical reasoning will enable us to determine local causal relationships between a target QoE variable, seen as effect, and multiple QoS metrics across or within levels, seen as causes. Such data-driven analysis is justified by the multiplicity of dependencies that exist between network or application QoS metrics as different adaptation mechanisms (e.g., TCP congestion avoidance, HTTP bitrate adaptation) are activated at each level in real life. Building optimal predictors based on (eventually several) probabilistically minimal subsets of features opens the way for a principled comparison of the predictors.

4.4. Crowd-sourced Information Filtering and Summarization

With the explosion of the People-centric Web, there is a proliferation of crowd-sourced content either under the form of qualitative reviews (mainly textual) and quantitative ratings (as 5 star ratings) regarding diverse products or services or under the form of various "real-time" feedback events (e.g., re-tweets, replies, likes, clicks, etc.) on published web content (ranging from traditional news, TV series, and movies to specialized blogs and posts shared over social networks). Such content captures the wisdom of the crowd and is valuable information source for building collaborative filtering systems and text summarization tools coping with information overload. For example, they can assist users to pick the most interesting web pages (e.g. Delicious) or to choose which movie to watch next (e.g. Netflix).

Implicit Feedback in Communities of a Place. We are initially interested in addressing one of the main limitation of collaborative filtering systems namely, the strong user engagement required to provide the necessary input (e.g., regarding their friends, tags or sites of preference) which is usual platform specific (i.e., for a particular social network, tagging, or bookmark system). The lack of user engagement translates into cold start and data sparsity. To cope with this limitation, we are developing a system called WeBrowse that passively observes network traffic to extract user clicks (i.e., the URLs users visit) for group of people who live, study, or work in the same place. Examples of such communities of a place are: (i) the students of a campus, (ii) the people living in a neighbourhood or (iii) researchers working in the same site. WeBrowse then promotes the hottest and most popular content to the community members sharing common interests.

Personalized Review Summarization. Finally, we are interested in helping people to take informed decisions regarding their shopping or entertainment activities. The automated summarization of a review corpus (for example, movie reviews from Rotten Tomatoes or IMDB; or restaurant reviews from Yelp) aims to assist people to form an opinion regarding a product/service of interest, by producing a coherent summary that is helpful and can be easily assimilated by humans. We are working on review summarisation methods that combine both objective (i.e., related to the review corpus) and subjective (i.e., related to the end-user interests) interestingness criteria of the produced reviews. In this respect we are exploiting domain models (e.g., Oscar's merit categories for movies) to elicit user preferences and mine the aspects of products/services actually commented in the textual sentences of reviews. For example, different summaries should be produced when a user is more interested in the actors' performance rather than the movie story. We are particularly interested in extracting automatically the signatures of aspects (based on a set of seed terms) and rank review sentences on their importance and relevance w.r.t. the aspects they comment. Last but not least we are optimizing the automatically constructed summary w.r.t. to a number of criteria such as the number of the length of included sentences from the original reviews, the polarity of sentiments in the described aspects, etc.

MYRIADS Project-Team

4. Application Domains

4.1. Main Application Domains

The Myriads team investigates the design and implementation of system services. Thus its research activities address a broad range of application domains. We validate our research results with selected use cases in the following application domains:

- Smart city services,
- Smart grids,
- Energy and sustainable development,
- Home IoT applications,
- Bio-informatics applications,
- Data science applications,
- Computational science applications,
- Numerical simulations.

REGAL Project-Team (section vide)

SPIRALS Project-Team

4. Application Domains

4.1. Introduction

Although our research is general enough to be applied to many application domains, we currently focus on applications and distributed services for the retail industry and for the digital home. These two application domains are supported by a strong expertise in mobile computing and in cloud computing that are the two main target environments on which our research prototypes are built, for which we are recognized, and for which we have already established strong collaborations with the industrial ecosystem.

4.2. Distributed software services for the retail industry

This application domain is developed in relation with the PICOM (*Pôle de compétivité Industries du Commerce*) cluster. We have established strong collaborations with local companies in the context of former funded projects, such as Cappucino and Macchiato, which focused on the development of a new generation of mobile computing platforms for e-commerce. We are also involved in the Datalyse and OCCIware funded projects that define cloud computing environments with applications for the retail industry. Finally, our activities in terms of crowd-sensing and data gathering on mobile devices with the APISENSE$^{®}$ platform share also applications for the retail industry.

4.3. Distributed software services for the digital home

We are developing new middleware solutions for the digital home, in particular through our long standing collaboration with Orange Labs. We are especially interested in developing energy management and saving solutions with the POWERAPI software library for distributed environments such the ones that equip digital homes. We are also working to bridge the gap between distributed services hosted on home gateways and distributed services hosted on the cloud to be able to smoothly transition between both environments. This work is especially conducted with the SALOON platform.

WHISPER Project-Team

4. Application Domains

4.1. Linux

Linux is an open-source operating system that is used in settings ranging from embedded systems to supercomputers. The most recent release of the Linux kernel, v4.14, comprises over 16 million lines of code, and supports 30 different families of CPU architectures, around 50 file systems, and thousands of device drivers. Linux is also in a rapid stage of development, with new versions being released roughly every 2.5 months. Recent versions have each incorporated around 13,500 commits, from around 1500 developers. These developers have a wide range of expertise, with some providing hundreds of patches per release, while others have contributed only one. Overall, the Linux kernel is critical software, but software in which the quality of the developed source code is highly variable. These features, combined with the fact that the Linux community is open to contributions and to the use of tools, make the Linux kernel an attractive target for software researchers. Tools that result from research can be directly integrated into the development of real software, where it can have a high, visible impact.

Starting from the work of Engler et al. [41], numerous research tools have been applied to the Linux kernel, typically for finding bugs [39], [58], [71], [81] or for computing software metrics [47], [87]. In our work, we have studied generic C bugs in Linux code [9], bugs in function protocol usage [52], [53], issues related to the processing of bug reports [75] and crash dumps [46], and the problem of backporting [70], [82], illustrating the variety of issues that can be explored on this code base. Unique among research groups working in this area, we have furthermore developed numerous contacts in the Linux developer community. These contacts provide insights into the problems actually faced by developers and serve as a means of validating the practical relevance of our work.

4.2. Device Drivers

Device drivers are essential to modern computing, to provide applications with access, via the operating system, to physical devices such as keyboards, disks, networks, and cameras. Development of new computing paradigms, such as the internet of things, is hampered because device driver development is challenging and error-prone, requiring a high level of expertise in both the targeted OS and the specific device. Furthermore, implementing just one driver is often not sufficient; today's computing landscape is characterized by a number of OSes, *e.g.*, Linux, Windows, MacOS, BSD and many real time OSes, and each is found in a wide range of variants and versions. All of these factors make the development, porting, backporting, and maintenance of device drivers a critical problem for device manufacturers, industry that requires specific devices, and even for ordinary users.

The last fifteen years have seen a number of approaches directed towards easing device driver development. Réveillère, who was supervised by G. Muller, proposes Devil [7], a domain-specific language for describing the low-level interface of a device. Chipounov *et al.* propose RevNic, [33] a template-based approach for porting device drivers from one OS to another. Ryzhyk *et al.* propose Termite, [72], [73] an approach for synthesizing device driver code from a specification of an OS and a device. Currently, these approaches have been successfully applied to only a small number of toy drivers. Indeed, Kadav and Swift [49] observe that these approaches make assumptions that are not satisfied by many drivers; for example, the assumption that a driver involves little computation other than the direct interaction between the OS and the device. At the same time, a number of tools have been developed for finding bugs in driver code. These tools include SDV [23], Coverity [41], CP-Miner, [57] PR-Miner [58], and Coccinelle [8]. These approaches, however, focus on analyzing existing code, and do not provide guidelines on structuring drivers.

In summary, there is still a need for a methodology that first helps the developer understand the software architecture of drivers for commonly used operating systems, and then provides tools for the maintenance of existing drivers.

<p style="text-align:center; color:red; font-weight:bold;">ALPINES Project-Team</p>

4. Application Domains

4.1. Compositional multiphase Darcy flow in heterogeneous porous media

We study the simulation of compositional multiphase flow in porous media with different types of applications, and we focus in particular on reservoir/bassin modeling, and geological CO2 underground storage. All these simulations are linearized using Newton approach, and at each time step and each Newton step, a linear system needs to be solved, which is the most expensive part of the simulation. This application leads to some of the difficult problems to be solved by iterative methods. This is because the linear systems arising in multiphase porous media flow simulations cumulate many difficulties. These systems are non-symmetric, involve several unknowns of different nature per grid cell, display strong or very strong heterogeneities and anisotropies, and change during the simulation. Many researchers focus on these simulations, and many innovative techniques for solving linear systems have been introduced while studying these simulations, as for example the nested factorization [Appleyard and Cheshire, 1983, SPE Symposium on Reservoir Simulation].

4.2. Inverse problems

The research of F. Nataf on inverse problems is rather new since this activity was started from scratch in 2007. Since then, several papers were published in international journals and conference proceedings. All our numerical simulations were performed in FreeFem++.

We focus on methods related to time reversal techniques. Since the seminal paper by [M. Fink et al., Imaging through inhomogeneous media using time reversal mirrors. Ultrasonic Imaging, 13(2):199, 1991.], time reversal is a subject of very active research. The main idea is to take advantage of the reversibility of wave propagation phenomena such as it occurs in acoustics, elasticity or electromagnetism in a non-dissipative unknown medium to back-propagate signals to the sources that emitted them. Number of industrial applications have already been developped: touchscreen, medical imaging, non-destructive testing and underwater communications. The principle is to back-propagate signals to the sources that emitted them. The initial experiment was to refocus, very precisely, a recorded signal after passing through a barrier consisting of randomly distributed metal rods. In [de Rosny and Fink. Overcoming the difraction limit in wave physics using a time-reversal mirror and a novel acoustic sink. Phys. Rev. Lett., 89 (12), 2002], the source that created the signal is time reversed in order to have a perfect time reversal experiment. Since then, numerous applications of this physical principle have been designed, see [Fink, Renversement du temps, ondes et innovation. Ed. Fayard, 2009] or for numerical experiments [Larmat et al., Time-reversal imaging of seismic sources and application to the great sumatra earthquake. Geophys. Res. Lett., 33, 2006] and references therein.

4.3. Numerical methods for wave propagation in multi-scale media

We are interested in the development of fast numerical methods for the simulation of electromagnetic waves in multi-scale situations where the geometry of the medium of propagation may be described through caracteristic lengths that are, in some places, much smaller than the average wavelength. In this context, we propose to develop numerical algorithms that rely on simplified models obtained by means of asymptotic analysis applied to the problem under consideration.

Here we focus on situations involving boundary layers and *localized* singular perturbation problems where wave propagation takes place in media whose geometry or material caracterics are submitted to a small scale perturbation localized around a point, or a surface, or a line, but not distributed over a volumic subregion of the propagation medium. Although a huge literature is already available for the study of localized singular perturbations and boundary layer pheneomena, very few works have proposed efficient numerical methods that rely on asymptotic modeling. This is due to their functional framework that naturally involves singular functions, which are difficult to handle numerically. The aim of this part of our reasearch is to develop and analyze numerical methods for singular perturbation methods that are prone to high order numerical approximation, and robust with respect to the small parameter characterizing the singular perturbation.

4.4. Data analysis in astrophysics

We focus on computationally intensive numerical algorithms arising in the data analysis of current and forthcoming Cosmic Microwave Background (CMB) experiments in astrophysics. This application is studied in collaboration with researchers from University Paris Diderot, and the objective is to make available the algorithms to the astrophysics community, so that they can be used in large experiments.

In CMB data analysis, astrophysicists produce and analyze multi-frequency 2D images of the universe when it was 5% of its current age. The new generation of the CMB experiments observes the sky with thousands of detectors over many years, producing overwhelmingly large and complex data sets, which nearly double every year therefore following Moore's Law. Planck (http://planck.esa.int/) is a keystone satellite mission which has been developed under auspices of the European Space Agency (ESA). Planck has been surveying the sky since 2010, produces terabytes of data and requires 100 Petaflops per image analysis of the universe. It is predicted that future experiments will collect half petabyte of data, and will require 100 Exaflops per analysis as early as in 2020. This shows that data analysis in this area, as many other applications, will keep pushing the limit of available supercomputing power for the years to come.

AVALON Project-Team

4. Application Domains

4.1. Overview

The Avalon team targets applications with large computing and/or data storage needs, which are still difficult to program, maintain, and deploy. Those applications can be parallel and/or distributed applications, such as large scale simulation applications or code coupling applications. Applications can also be workflow-based as commonly found in distributed systems such as grids or clouds.

The team aims at not being restricted to a particular application field, thus avoiding any spotlight. The team targets different HPC and distributed application fields, which bring use cases with different issues. This will be eased by our various collaborations: the team participates to the INRIA-Illinois Joint Laboratory for Petascale Computing, the Physics, Radiobiology, Medical Imaging, and Simulation French laboratory of excellence, the E-Biothon project, the INRIA large scale initiative Computer and Computational Sciences at Exascale (C2S@Exa), and to BioSyL, a federative research structure about Systems Biology of the University of Lyon. Moreover, the team members have a long tradition of cooperation with application developers such as CERFACS and EDF R&D. Last but not least, the team has a privileged connection with CC IN2P3 that opens up collaborations, in particular in the astrophysics field.

In the following, some examples of representative applications we are targeting are presented. In addition to highlighting some application needs, they also constitute some of the use cases we will use to valide our theoretical results.

4.2. Climatology

The world's climate is currently changing due to the increase of the greenhouse gases in the atmosphere. Climate fluctuations are forecasted for the years to come. For a proper study of the incoming changes, numerical simulations are needed, using general circulation models of a climate system. Simulations can be of different types: HPC applications (*e.g.,* the NEMO framework [38] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [44] for global climate modeling), or workflows (long term global climate modeling).

As for most applications the team is targeting, the challenge is to thoroughly analyze climate-forecasting applications to model their needs in terms of programing model, execution model, energy consumption, data access pattern, and computing needs. Once a proper model of an application has been set up, appropriate scheduling heuristics could be designed, tested, and compared. The team has a long tradition of working with CERFACS on this topic, for example in the LEGO (2006-09) and SPADES (2009-12) French ANR projects.

4.3. Astrophysics

Astrophysics is a major field to produce large volumes of data. For instance, the Large Synoptic Survey Telescope (http://www.lsst.org/lsst/) will produce 15 TB of data every night, with the goals of discovering thousands of exoplanets and of uncovering the nature of dark matter and dark energy in the universe. The Square Kilometer Array (http://www.skatelescope.org/) produces 9 Tbits/s of raw data. One of the scientific projects related to this instrument called Evolutionary Map of the Universe is working on more than 100 TB of images. The Euclid Imaging Consortium (https://www.euclid-ec.org/) will generate 1 PB data per year.

Avalon collaborates with the *Institut de Physique Nucléaire de Lyon* (IPNL) laboratory on large scale numerical simulations in astronomy and astrophysics. Contributions of the Avalon members have been related to algorithmic skeletons to demonstrate large scale connectivity, the development of procedures for the generation of realistic mock catalogs, and the development of a web interface to launch large cosmological simulations on GRID'5000.

This collaboration, that continues around the topics addressed by the CLUES project (http://www.clues-project.org), has been extended thanks to the tight links with the CC-IN2P3. Major astrophysics projects execute part of their computing, and store part of their data on the resources provided by the CC-IN2P3. Among them, we can mention SNFactory, Euclid, or LSST. These applications constitute typical use cases for the research developed in the Avalon team: they are generally structured as workflows and a huge amount of data (from TB to PB) is involved.

4.4. Bioinformatics

Large-scale data management is certainly one of the most important applications of distributed systems in the future. Bioinformatics is a field producing such kinds of applications. For example, DNA sequencing applications make use of MapReduce skeletons.

The Avalon team is a member of BioSyL (http://www.biosyl.org), a Federative Research Structure attached to University of Lyon. It gathers about 50 local research teams working on systems biology. Moreover, the team cooperated with the French Institute of Biology and Chemistry of Proteins (IBCP http://www.ibcp.fr) in particular through the ANR MapReduce project where the team focuses on a bio-chemistry application dealing with protein structure analysis. Avalon has also started working with the Inria Beagle team (https://team.inria.fr/beagle/) on artificial evolution and computational biology as the challenges are around high performance computation and data management.

<h2 style="color:red; text-align:center">DATAMOVE Project-Team</h2>

4. Application Domains

4.1. Data Aware Batch Scheduling

Large scale high performance computing platforms are becoming increasingly complex. Determining efficient allocation and scheduling strategies that can adapt to technological evolutions is a strategic and difficult challenge. We are interested in scheduling jobs in hierarchical and heterogeneous large scale platforms. On such platforms, application developers typically submit their jobs in centralized waiting queues. The job management system aims at determining a suitable allocation for the jobs, which all compete against each other for the available computing resources. Performances are measured using different classical metrics like maximum completion time or slowdown. Current systems make use of very simple (but fast) algorithms that however rely on simplistic platform and execution models, and thus, have limited performances.

For all target scheduling problems we aim to provide both theoretical analysis and complementary analysis through simulations. Achieving meaningful results will require strong improvements on existing models (on power for example) and the design of new approximation algorithms with various objectives such as stretch, reliability, throughput or energy consumption, while keeping in focus the need for a low-degree polynomial complexity.

4.1.1. Algorithms

The most common batch scheduling policy is to consider the jobs according to the First Come First Served order (FCFS) with backfilling (BF). BF is the most widely used policy due to its easy and robust implementation and known benefits such as high system utilization. It is well-known that this strategy does not optimize any sophisticated function, but it is simple to implement and it guarantees that there is no starvation (i.e. every job will be scheduled at some moment).

More advanced algorithms are seldom used on production platforms due to both the gap between theoretical models and practical systems and speed constraints. When looking at theoretical scheduling problems, the generally accepted goal is to provide polynomial algorithms (in the number of submitted jobs and the number of involved computing units). However, with millions of processing cores where every process and data transfer have to be individually scheduled, polynomial algorithms are prohibitive as soon as the polynomial degree is too large. The model of *parallel tasks* simplifies this problem by bundling many threads and communications into single boxes, either rigid, rectangular or malleable. Especially malleable tasks capture the dynamicity of the execution. Yet these models are ill-adapted to heterogeneous platforms, as the running time depends on more than simply the number of allotted resources, and some of the common underlying assumptions on the speed-up functions (such as monotony or concavity) are most often only partially verified.

In practice, the job execution times depend on their allocation (due to communication interferences and heterogeneity in both computation and communication), while theoretical models of parallel jobs usually consider jobs as black boxes with a fixed (maximum) execution time. Though interesting and powerful, the classical models (namely, synchronous PRAM model, delay, LogP) and their variants (such as hierarchical delay), are not well-suited to large scale parallelism on platforms where the cost of moving data is significant, non uniform and may change over time. Recent studies are still refining such models in order to take into account communication contentions more accurately while remaining tractable enough to provide a useful tool for algorithm design.

Today, all algorithms in use in production systems are oblivious to communications. One of our main goals is to **design a new generation of scheduling algorithms fitting more closely job schedules according to platform topologies**.

4.1.2. *Locality Aware Allocations*

Recently, we developed modifications of the standard back-filling algorithm taking into account platform topologies. The proposed algorithms take into account locality and contiguity in order to hide communication patterns within parallel tasks. The main result here is to establish good lower bounds and small approximation ratios for policies respecting the locality constraints. The algorithms work in an online fashion, improving the global behavior of the system while still keeping a low running time. These improvements rely mainly on our past experience in designing approximation algorithms. Instead of relying on complex networking models and communication patterns for estimating execution times, the communications are disconnected from the execution time. Then, the scheduling problem leads to a trade-off: optimizing locality of communications on one side and a performance objective (like the makespan or stretch) on the other side.

In the perspective of taking care of locality, other ongoing works include the study of schedulers for platforms whose interconnection network is a static structured topology (like the 3D-torus of the BlueWaters platform we work on in collaboration with the Argonne National Laboratory). One main characteristic of this 3D-torus platform is to provide I/O nodes at specific locations in the topology. Applications generate and access specific data and are thus bounded to specific I/O nodes. Resource allocations are constrained in a strong and unusual way. This problem is close for actual hierarchical platforms. The scheduler needs to compute a schedule such that I/O nodes requirements are filled for each application while at the same time avoiding communication interferences. Moreover, extra constraints can arise for applications requiring accelerators that are gathered on the nodes at the edge of the network topology.

While current results are encouraging, they are however limited in performance by the low amount of information available to the scheduler. We look forward to extend ongoing work by progressively increasing application and network knowledge (by technical mechanisms like profiling or monitoring or by more sophisticated methods like learning). It is also important to anticipate on application resource usage in terms of compute units, memory as well as network and I/Os to efficiently schedule a mix of applications with different profiles. For instance, a simple solution is to partition the jobs as "communication intensive" or "low communications". Such a tag could be achieved by the users them selves or obtained by learning techniques. We could then schedule low communications jobs using leftover spaces while taking care of high communication jobs. More sophisticated options are possible, for instance those that use more detailed communication patterns and networking models. Such options would leverage the work proposed in Section 4.2 for gathering application traces.

4.1.3. *Data-Centric Processing*

Exascale computing is shifting away from the traditional compute-centric models to a more data-centric one. This is driven by the evolving nature of large scale distributed computing, no longer dominated by pure computations but also by the need to handle and analyze large volumes of data. These data can be large databases of results, data streamed from a running application or another scientific instrument (collider for instance). These new workloads call for specific resource allocation strategies.

Data movements and storage are expected to be a major energy and performance bottleneck on next generation platforms. Storage architectures are also evolving, the standard centralized parallel file system being complemented with local persistent storage (Burst Buffers, NVRAM). Thus, one data producer can stage data on some nodes' local storage, requiring to schedule close by the associated analytics tasks to limit data movements. This kind of configuration, often referred as *in-situ analytics*, is expected to become common as it enables to switch from the traditional I/O intensive workflow (batch-processing followed by *post mortem* analysis and visualization) to a more storage conscious approach where data are processed as closely as possible to where and when they are produced (in-situ processing is addressed in details in section 4.3). By reducing data movements and scheduling the extra processing on resources not fully exploited yet, in-situ processing is expected to have also a significant positive energetic impact. Analytics codes can be executed in the same nodes than the application, often on dedicated cores commonly called helper cores, or on dedicated nodes called staging nodes. The results are either forwarded to the users for visualization or saved to disk through I/O nodes. In-situ analytics can also take benefit of node local disks or burst buffers to reduce data movements.

Future job scheduling strategies should take into account in-situ processes in addition to the job allocation to optimize both energy consumption and execution time. On the one hand, this problem can be reduced to an allocation problem of extra asynchronous tasks to idle computing units. But on the other hand, embedding analytics in applications brings extra difficulties by making the application more heterogeneous and imposing more constraints (data affinity) on the required resources. Thus, the main point here is to develop efficient algorithms for dealing with heterogeneity without increasing the global computational cost.

4.1.4. Learning

Another important issue is to adapt the job management system to deal with the bad effects of uncertainties, which may be catastrophic in large scale heterogeneous HPC platforms (jobs delayed arbitrarly far or jobs killed). A natural question is then: *is it possible to have a good estimation of the job and platform parameters in order to be able to obtain a better scheduling ?* Many important parameters (like the number or type of required resources or the estimated running time of the jobs) are asked to the users when they submit their jobs. However, some of these values are not accurate and in many cases, they are not even provided by the end-users. In DataMove, we propose to study new methods for a better prediction of the characteristics of the jobs and their execution in order to improve the optimization process. In particular, the methods well-studied in the field of big data (in supervised Machine Learning, like classical regression methods, Support Vector Methods, random forests, learning to rank techniques or deep learning) could and must be used to improve job scheduling in large scale HPC platforms. This topic received a great attention recently in the field of parallel and distributed processing. A preliminary study has been done recently by our team with the target of predicting the job running times (called wall times). We succeeded to improve significantly in average the reference EASY Back Filling algorithm by estimating the wall time of the jobs, however, this method leads to big delay for the stretch of few jobs. Even if we succeed in determining more precisely hidden parameters, like the wall time of the jobs, this is not enough to determine an optimized solution. The shift is not only to learn on dedicated parameters but also on the scheduling policy. The data collected from the accounting and profiling of jobs can be used to better understand the needs of the jobs and through learning to propose adaptations for future submissions. The goal is to propose extensions to further improve the job scheduling and improve the performance and energy efficiency of the application. For instance preference learning may enable to compute on-line new priorities to back-fill the ready jobs.

4.1.5. Multi-objective Optimization

Several optimization questions that arise in allocation and scheduling problems lead to the study of several objectives at the same time. The goal is then not a single optimal solution, but a more complicated mathematical object that captures the notion of trade-off. In broader terms, the goal of multi-objective optimization is not to externally arbitrate on disputes between entities with different goals, but rather to explore the possible solutions to highlight the whole range of interesting compromises. A classical tool for studying such multi-objective optimization problems is to use *Pareto curves*. However, the full description of the Pareto curve can be very hard because of both the number of solutions and the hardness of computing each point. Addressing this problem will opens new methodologies for the analysis of algorithms.

To further illustrate this point here are three possible case studies with emphasis on conflicting interests measured with different objectives. While these cases are good representatives of our HPC context, there are other pertinent trade-offs we may investigate depending on the technology evolution in the coming years. This enumeration is certainly not limitative.

Energy versus Performance. The classical scheduling algorithms designed for the purpose of performance can no longer be used because performance and energy are contradictory objectives to some extent. The scheduling problem with energy becomes a multi-objective problem in nature since the energy consumption should be considered as equally important as performance at exascale. A global constraint on energy could be a first idea for determining trade-offs but the knowledge of the Pareto set (or an approximation of it) is also very useful.

Administrators versus application developers. Both are naturally interested in different objectives: In current algorithms, the performance is mainly computed from the point of view of administrators, but the users should be in the loop since they can give useful information and help to the construction of better schedules. Hence, we face again a multi-objective problem where, as in the above case, the approximation of the Pareto set provides the trade-off between the administrator view and user demands. Moreover, the objectives are usually of the same nature. For example, *max stretch* and *average stretch* are two objectives based on the slowdown factor that can interest administrators and users, respectively. In this case the study of the norm of stretch can be also used to describe the trade-off (recall that the L_1-norm corresponds to the average objective while the L_∞-norm to the max objective). Ideally, we would like to design an algorithm that gives good approximate solutions at the same time for all norms. The L_2 or L_3-norm are useful since they describe the performance of the whole schedule from the administrator point of view as well as they provide a fairness indication to the users. The hard point here is to derive theoretical analysis for such complicated tools.

Resource Augmentation. The classical resource augmentation models, i.e. speed and machine augmentation, are not sufficient to get good results when the execution of jobs cannot be frequently interrupted. However, based on a resource augmentation model recently introduced, where the algorithm may reject a small number of jobs, some members of our team have given the first interesting results in the non-preemptive direction. In general, resource augmentation can explain the intuitive good behavior of some greedy algorithms while, more interestingly, it can give ideas for new algorithms. For example, in the rejection context we could dedicate a small number of nodes for the usually problematic rejected jobs. Some initial experiments show that this can lead to a schedule for the remaining jobs that is very close to the optimal one.

4.2. Empirical Studies of Large Scale Platforms

Experiments or realistic simulations are required to take into account the impact of allocations and assess the real behavior of scheduling algorithms. While theoretical models still have their interest to lay the groundwork for algorithmic designs, the models are necessarily reflecting a purified view of the reality. As transferring our algorithm in a more practical setting is an important part of our creed, we need to ensure that the theoretical results found using simplified models can really be transposed to real situations. On the way to exascale computing, large scale systems become harder to study, to develop or to calibrate because of the costs in both time and energy of such processes. It is often impossible to convince managers to use a production cluster for several hours simply to test modifications in the RJMS. Moreover, as the existing RJMS production systems need to be highly reliable, each evolution requires several real scale test iterations. The consequence is that scheduling algorithms used in production systems are mostly outdated and not customized correctly. To circumvent this pitfall, we need to develop tools and methodologies for alternative empirical studies, from analysis of workload traces, to job models, simulation and emulation with reproducibility concerns.

4.2.1. Workload Traces with Resource Consumption

Workload traces are the base element to capture the behavior of complete systems composed of submitted jobs, running applications, and operating tools. These traces must be obtained on production platforms to provide relevant and representative data. To get a better understanding of the use of such systems, we need to look at both, how the jobs interact with the job management system, and how they use the allocated resources. We propose a general workload trace format that adds jobs resource consumption to the commonly used SWF [0] workload trace format. This requires to instrument the platforms, in particular to trace resource consumptions like CPU, data movements at memory, network and I/O levels, with an acceptable performance impact. In a previous work we studied and proposed a dedicated job monitoring tool whose impact on the system has been measured as lightweight (0.35% speed-down) with a 1 minute sampling rate. Other tools also explore job monitoring, like TACC Stats. A unique feature from our tool is its ability to monitor distinctly jobs sharing common nodes.

---

[0]Standard Workload Format: http://www.cs.huji.ac.il/labs/parallel/workload/swf.html

Collected workload traces with jobs resource consumption will be publicly released and serve to provide data for works presented in Section 4.1 . The trace analysis is expected to give valuable insights to define models encompassing complex behaviours like network topology sensitivity, network congestion and resource interferences.

We expect to join efforts with partners for collecting quality traces (ATOS/Bull, Ciment meso center, Joint Laboratory on Extreme Scale Computing) and will collaborate with the Inria team POLARIS for their analysis.

4.2.2. *Simulation*

Simulations of large scale systems are faster by multiple orders of magnitude than real experiments. Unfortunately, replacing experiments with simulations is not as easy as it may sound, as it brings a host of new problems to address in order to ensure that the simulations are closely approximating the execution of typical workloads on real production clusters. Most of these problems are actually not directly related to scheduling algorithms assessment, in the sense that the workload and platform models should be defined independently from the algorithm evaluations, in order to ensure a fair assessment of the algorithms' strengths and weaknesses. These research topics (namely platform modeling, job models and simulator calibration) are addressed in the other subsections.

We developed an open source platform simulator within DataMove (in conjunction with the OAR development team) to provide a widely distributable test bed for reproducible scheduling algorithm evaluation. Our simulator, named Batsim, allows to simulate the behavior of a computational platform executing a workload scheduled by any given scheduling algorithm. To obtain sound simulation results and to broaden the scope of the experiments that can be done thanks to Batsim, we did not chose to create a (necessarily limited) simulator from scratch, but instead to build on top of the SimGrid simulation framework.

To be open to as many batch schedulers as possible, Batsim decouples the platform simulation and the scheduling decisions in two clearly-separated software components communicating through a complete and documented protocol. The Batsim component is in charge of simulating the computational resources behaviour whereas the scheduler component is in charge of taking scheduling decisions. The scheduler component may be both a resource and a job management system. For jobs, scheduling decisions can be to execute a job, to delay its execution or simply to reject it. For resources, other decisions can be taken, for example to change the power state of a machine i.e. to change its speed (in order to lower its energy consumption) or to switch it on or off. This separation of concerns also enables interfacing with potentially any commercial RJMS, as long as the communication protocol with Batsim is implemented. A proof of concept is already available with the OAR RJMS.

Using this test bed opens new research perspectives. It allows to test a large range of platforms and workloads to better understand the real behavior of our algorithms in a production setting. In turn, this opens the possibility to tailor algorithms for a particular platform or application, and to precisely identify the possible shortcomings of the theoretical models used.

4.2.3. *Job and Platform Models*

The central purpose of the Batsim simulator is to simulate job behaviors on a given target platform under a given resource allocation policy. Depending on the workload, a significant number of jobs are parallel applications with communications and file system accesses. It is not conceivable to simulate individually all these operations for each job on large plaforms with their associated workload due to implied simulation complexity. The challenge is to define a coarse grain job model accurate enough to reproduce parallel application behavior according to the target platform characteristics. We will explore models similar to the BSP (Bulk Synchronous Program) approach that decomposes an application in local computation supersteps ended by global communications and a global synchronization. The model parameters will be established by means of trace analysis as discussed previously, but also by instrumenting some parallel applications to capture communication patterns. This instrumentation will have a significant impact on the concerned application performance, restricting its use to a few applications only. There are a lot of recurrent applications executed on HPC platform, this fact will help to reduce the required number of instrumentations and captures. To assign

each job a model, we are considering to adapt the concept of application signatures as proposed in. Platform models and their calibration are also required. Large parts of these models, like those related to network, are provided by Simgrid. Other parts as the filesystem and energy models are comparatively recent and will need to be enhanced or reworked to reflect the HPC platform evolutions. These models are then generally calibrated by running suitable benchmarks.

4.2.4. Emulation and Reproducibility

The use of coarse models in simulation implies to set aside some details. This simplification may hide system behaviors that could impact significantly and negatively the metrics we try to enhance. This issue is particularly relevant when large scale platforms are considered due to the impossibility to run tests at nominal scale on these real platforms. A common approach to circumvent this issue is the use of emulation techniques to reproduce, under certain conditions, the behavior of large platforms on smaller ones. Emulation represents a natural complement to simulation by allowing to execute directly large parts of the actual evaluated software and system, but at the price of larger compute times and a need for more resources. The emulation approach was chosen in to compare two job management systems from workload traces of the CURIE supercomputer (80000 cores). The challenge is to design methods and tools to emulate with sufficient accuracy the platform and the workload (data movement, I/O transfers, communication, applications interference). We will also intend to leverage emulation tools like Distem from the MADYNES team. It is also important to note that the Batsim simulator also uses emulation techniques to support the core scheduling module from actual RJMS. But the integration level is not the same when considering emulation for larger parts of the system (RJMS, compute node, network and filesystem).

Replaying traces implies to prepare and manage complex software stacks including the OS, the resource management system, the distributed filesystem and the applications as well as the tools required to conduct experiments. Preparing these stacks generate specific issues, one of the major one being the support for reproducibility. We propose to further develop the concept of reconstructability to improve experiment reproducibility by capturing the build process of the complete software stack. This approach ensures reproducibility over time better than other ways by keeping all data (original packages, build recipe and Kameleon engine) needed to build the software stack.

In this context, the Grid'5000 (see Sec. 6.4) experimentation infrastructure that gives users the control on the complete software stack is a crucial tool for our research goals. We will pursue our strong implication in this infrastructure.

4.3. Integration of High Performance Computing and Data Analytics

Data produced by large simulations are traditionally handled by an I/O layer that moves them from the compute cores to the file system. Analysis of these data are performed after reading them back from files, using some domain specific codes or some scientific visualisation libraries like VTK. But writing and then reading back these data generates a lot of data movements and puts under pressure the file system. To reduce these data movements, **the in situ analytics paradigm proposes to process the data as closely as possible to where and when the data are produced**. Some early solutions emerged either as extensions of visualisation tools or of I/O libraries like ADIOS. But significant progresses are still required to provide efficient and flexible high performance scientific data analysis tools. Integrating data analytics in the HPC context will have an impact on resource allocation strategies, analysis algorithms, data storage and access, as well as computer architectures and software infrastructures. But this paradigm shift imposed by the machine performance also sets the basis for a deep change on the way users work with numerical simulations. The traditional workflow needs to be reinvented to make HPC more user-centric, more interactive and turn HPC into a commodity tool for scientific discovery and engineering developments. In this context DataMove aims at investigating programming environments for in situ analytics with a specific focus on task scheduling in particular, to ensure an efficient sharing of resources with the simulation.

4.3.1. Programming Model and Software Architecture

In situ creates a tighter loop between the scientist and her/his simulation. As such, an in situ framework needs to be flexible to let the user define and deploy its own set of analysis. A manageable flexibility requires to favor simplicity and understandability, while still enabling an efficient use of parallel resources. Visualization libraries like VTK or Visit, as well as domain specific environments like VMD have initially been developed for traditional post-mortem data analysis. They have been extended to support in situ processing with some simple resource allocation strategies but the level of performance, flexibility and ease of use that is expected requires to rethink new environments. There is a need to develop a middleware and programming environment taking into account in its fundations this specific context of high performance scientific analytics.

Similar needs for new data processing architectures occurred for the emerging area of Big Data Analytics, mainly targeted to web data on cloud-based infrastructures. Google Map/Reduce and its successors like Spark or Stratosphere/Flink have been designed to match the specific context of efficient analytics for large volumes of data produced on the web, on social networks, or generated by business applications. These systems have mainly been developed for cloud infrastructures based on commodity architectures. They do not leverage the specifics of HPC infrastructures. Some preliminary adaptations have been proposed for handling scientific data in a HPC context. However, these approaches do not support in situ processing.

Following the initial development of FlowVR, our middleware for in situ processing, we will pursue our effort to develop a programming environment and software architecture for high performance scientific data analytics. Like FlowVR, the map/reduce tools, as well as the machine learning frameworks like TensorFlow, adopted a dataflow graph for expressing analytics pipe-lines. We are convinced that this dataflow approach is both easy to understand and yet expresses enough concurrency to enable efficient executions. The graph description can be compiled towards lower level representations, a mechanism that is intensively used by Stratosphere/Flink for instance. Existing in situ frameworks, including FlowVR, inherit from the HPC way of programming with a thiner software stack and a programming model close to the machine. Though this approach enables to program high performance applications, this is usually too low level to enable the scientist to write its analysis pipe-line in a short amount of time. The data model, i.e. the data semantics level accessible at the framework level for error check and optimizations, is also a fundamental aspect of such environments. The key/value store has been adopted by all map/reduce tools. Except in some situations, it cannot be adopted as such for scientific data. Results from numerical simulations are often more structured than web data, associated with acceleration data structures to be processed efficiently. We will investigate data models for scientific data building on existing approaches like Adios or DataSpaces.

4.3.2. Resource Sharing

To alleviate the I/O bottleneck, the in situ paradigm proposes to start processing data as soon as made available by the simulation, while still residing in the memory of the compute node. In situ processings include data compression, indexing, computation of various types of descriptors (1D, 2D, images, etc.). Per se, reducing data output to limit I/O related performance drops or keep the output data size manageable is not new. Scientists have relied on solutions as simple as decreasing the frequency of result savings. In situ processing proposes to move one step further, by providing a full fledged processing framework enabling scientists to more easily and thoroughly manage the available I/O budget.

The most direct way to perform in situ analytics is to inline computations directly in the simulation code. In this case, in situ processing is executed in sequence with the simulation that is suspended meanwhile. Though this approach is direct to implement and does not require complex framework environments, it does not enable to overlap analytics related computations and data movements with the simulation execution, preventing to efficiently use the available resources. Instead of relying on this simple time sharing approach, several works propose to rely on space sharing where one or several cores per node, called *helper cores*, are dedicated to analytics. The simulation responsibility is simply to handle a copy of the relevant data to the node-local in situ processes, both codes being executed concurrently. This approach often lead to significantly beter performance than in-simulation analytics.

For a better isolation of the simulation and in situ processes, one solution consists in offloading in situ tasks from the simulation nodes towards extra dedicated nodes, usually called *staging nodes*. These computations are said to be performed *in-transit*. But this approach may not always be beneficial compared to processing on simulation nodes due to the costs of moving the data from the simulation nodes to the staging nodes.

FlowVR enables to mix these different resources allocation strategies for the different stages of an analytics pile-line. Based on a component model, the scientist designs analytics workflows by first developing processing components that are next assembled in a dataflow graph through a Python script. At runtime the graph is instantiated according to the execution context, FlowVR taking care of deploying the application on the target architecture, and of coordinating the analytics workflows with the simulation execution.

But today the choice of the resource allocation strategy is mostly ad-hoc and defined by the programmer. We will investigate solutions that enable a cooperative use of the resource between the analytics and the simulation with minimal hints from the programmer. In situ processings inherit from the parallelization scale and data distribution adopted by the simulation, and must execute with minimal perturbations on the simulation execution (whose actual resource usage is difficult to know a priori). We need to develop adapted scheduling strategies that operate at compile and run time. Because analysis are often data intensive, such solutions must take into consideration data movements, a point that classical scheduling strategies designed first for compute intensive applications often overlook. We expect to develop new scheduling strategies relying on the methodologies developed in Sec. 4.1.5 . Simulations as well as analysis are iterative processes exposing a strong spatial and temporal coherency that we can take benefit of to anticipate their behavior and then take more relevant resources allocation strategies, possibly based on advanced learning algorithms or as developed in Section 4.1 .

In situ analytics represent a specific workload that needs to be scheduled very closely to the simulation, but not necessarily active during the full extent of the simulation execution and that may also require to access data from previous runs (stored in the file system or on specific burst-buffers). Several users may also need to run concurrent analytics pipe-lines on shared data. This departs significantly from the traditional batch scheduling model, motivating the need for a more elastic approach to resource provisioning. These issues will be conjointly addressed with research on batch scheduling policies (Sec. 4.1).

4.3.3. *Co-Design with Data Scientists*

Given the importance of users in this context, it is of primary importance that in situ tools be co-designed with advanced users, even if such multidisciplinary collaborations are challenging and require constant long term investments to learn and understand the specific practices and expectations of the other domain.

We will tightly collaborate with scientists of some application domains, like molecular dynamics or fluid simulation, to design, develop, deploy and assess in situ analytics scenarios, as already done with Marc Baaden, a computational biologist from LBT.

We recently extended our collaboration network. We started in 2015 a PhD co-advised with CEA DAM to investigate in situ analytics scenarios in the context of atomistic material simulations. CEA DAM is a French energy lab hosting one of the largest european supercomputer. They gather physicists, numerical scientists as well as high performance computer engineers, making it a very interesting partner for developing new scientific data analysis solutions. We also got a national grant (2015-2018) to compute in situ statistics for multi-parametric parallel studies with the research department of French power company EDF. In this context we collaborate with statisticians and fluid simulation experts to define in situ scenarios, revisit the statistic operators to be amenable to in situ processing, and define an adapted in situ framework.

HIEPACS Project-Team

4. Application Domains

4.1. Material physics

Participants: Pierre Blanchard, Olivier Coulaud, Arnaud Durocher.

Due to the increase of available computer power, new applications in nano science and physics appear such as study of properties of new materials (photovoltaic materials, bio- and environmental sensors, ...), failure in materials, nano-indentation. Chemists, physicists now commonly perform simulations in these fields. These computations simulate systems up to billion of atoms in materials, for large time scales up to several nanoseconds. The larger the simulation, the smaller the computational cost of the potential driving the phenomena, resulting in low precision results. So, if we need to increase the precision, there are two ways to decrease the computational cost. In the first approach, we improve algorithms and their parallelization and in the second way, we will consider a multiscale approach.

A domain of interest is the material aging for the nuclear industry. The materials are exposed to complex conditions due to the combination of thermo-mechanical loading, the effects of irradiation and the harsh operating environment. This operating regime makes experimentation extremely difficult and we must rely on multi-physics and multi-scale modeling for our understanding of how these materials behave in service. This fundamental understanding helps not only to ensure the longevity of existing nuclear reactors, but also to guide the development of new materials for 4th generation reactor programs and dedicated fusion reactors. For the study of crystalline materials, an important tool is dislocation dynamics (DD) modeling. This multiscale simulation method predicts the plastic response of a material from the underlying physics of dislocation motion. DD serves as a crucial link between the scale of molecular dynamics and macroscopic methods based on finite elements; it can be used to accurately describe the interactions of a small handful of dislocations, or equally well to investigate the global behavior of a massive collection of interacting defects.

To explore i.e. to simulate these new areas, we need to develop and/or to improve significantly models, schemes and solvers used in the classical codes. In the project, we want to accelerate algorithms arising in those fields. We will focus on the following topics (in particular in the currently under definition OPTIDIS project in collaboration with CEA Saclay, CEA Ile-de-france and SIMaP Laboratory in Grenoble) in connection with research described at Sections 3.4 and 3.5 .

- The interaction between dislocations is long ranged ($O(1/r)$) and anisotropic, leading to severe computational challenges for large-scale simulations. In dislocation codes, the computation of interaction forces between dislocations is still the most CPU time consuming and has to be improved to obtain faster and more accurate simulations.

- In such simulations, the number of dislocations grows while the phenomenon occurs and these dislocations are not uniformly distributed in the domain. This means that strategies to dynamically construct a good load balancing are crucial to acheive high performance.

- From a physical and a simulation point of view, it will be interesting to couple a molecular dynamics model (atomistic model) with a dislocation one (mesoscale model). In such three-dimensional coupling, the main difficulties are firstly to find and characterize a dislocation in the atomistic region, secondly to understand how we can transmit with consistency the information between the two micro and meso scales.

4.2. Co-design for scalable numerical algorithms in scientific applications

Participants: Nicolas Bouzat, Mathieu Faverge, Guillaume Latu, Michel Mehrenberger, Pierre Ramet, Jean Roman.

4.2.1. High performance simulation for ITER tokamak

Scientific simulation for ITER tokamak modeling provides a natural bridge between theory and experimentation and is also an essential tool for understanding and predicting plasma behavior. Recent progresses in numerical simulation of fine-scale turbulence and in large-scale dynamics of magnetically confined plasma have been enabled by access to petascale supercomputers. These progresses would have been unreachable without new computational methods and adapted reduced models. In particular, the plasma science community has developed codes for which computer runtime scales quite well with the number of processors up to thousands cores. The research activities of HIEPACS concerning the international ITER challenge were involved in the Inria Project Lab C2S@EXA in collaboration with CEA-IRFM and are related to two complementary studies: a first one concerning the turbulence of plasma particles inside a tokamak (in the context of GYSELA code) and a second one concerning the MHD instability edge localized modes (in the context of JOREK code).

Currently, GYSELA is parallelized in an hybrid MPI+OpenMP way and can exploit the power of the current greatest supercomputers. To simulate faithfully the plasma physic, GYSELA handles a huge amount of data and today, the memory consumption is a bottleneck on very large simulations. In this context, mastering the memory consumption of the code becomes critical to consolidate its scalability and to enable the implementation of new numerical and physical features to fully benefit from the extreme scale architectures.

Other numerical simulation tools designed for the ITER challenge aim at making a significant progress in understanding active control methods of plasma edge MHD instability Edge Localized Modes (ELMs) which represent a particular danger with respect to heat and particle loads for Plasma Facing Components (PFC) in the tokamak. The goal is to improve the understanding of the related physics and to propose possible new strategies to improve effectiveness of ELM control techniques. The simulation tool used (JOREK code) is related to non linear MHD modeling and is based on a fully implicit time evolution scheme that leads to 3D large very badly conditioned sparse linear systems to be solved at every time step. In this context, the use of PaStiX library to solve efficiently these large sparse problems by a direct method is a challenging issue.

KERDATA Project-Team (section vide)

<p style="text-align:center">**POLARIS Team**</p>

4. Application Domains

4.1. Large Computing Infrastructures

Supercomputers typically comprise thousands to millions of multi-core CPUs with GPU accelerators interconnected by complex interconnection networks that are typically structured as an intricate hierarchy of network switches. Capacity planning and management of such systems not only raises challenges in term of computing efficiency but also in term of energy consumption. Most legacy (SPMD) applications struggle to benefit from such infrastructure since the slightest failure or load imbalance immediately causes the whole program to stop or at best to waste resources. To scale and handle the stochastic nature of resources, these applications have to rely on dynamic runtimes that schedule computations and communications in an opportunistic way. Such evolution raises challenges not only in terms of programming but also in terms of observation (complexity and dynamicity prevents experiment reproducibility, intrusiveness hinders large scale data collection, ...) and analysis (dynamic and flexible application structures make classical visualization and simulation techniques totally ineffective and require to build on *ad hoc* information on the application structure).

4.2. Next-Generation Wireless Networks

Considerable interest has arisen from the seminal prediction that the use of multiple-input, multiple-output (MIMO) technologies can lead to substantial gains in information throughput in wireless communications, especially when used at a massive level. In particular, by employing multiple inexpensive service antennas, it is possible to exploit spatial multiplexing in the transmission and reception of radio signals, the only physical limit being the number of antennas that can be deployed on a portable device. As a result, the wireless medium can accommodate greater volumes of data traffic without requiring the reallocation (and subsequent re-regulation) of additional frequency bands. In this context, throughput maximization in the presence of interference by neighboring transmitters leads to games with convex action sets (covariance matrices with trace constraints) and individually concave utility functions (each user's Shannon throughput); developing efficient and distributed optimization protocols for such systems is one of the core objectives of Theme 5.

Another major challenge that occurs here is due to the fact that the efficient physical layer optimization of wireless networks relies on perfect (or close to perfect) channel state information (CSI), on both the uplink and the downlink. Due to the vastly increased computational overhead of this feedback – especially in decentralized, small-cell environments – the ongoing transition to fifth generation (5G) wireless networks is expected to go hand-in-hand with distributed learning and optimization methods that can operate reliably in feedback-starved environments. Accordingly, one of POLARIS' application-driven goals will be to leverage the algorithmic output of Theme 5 into a highly adaptive resource allocation framework for next-gneration wireless systems that can effectively "learn in the dark", without requiring crippling amounts of feedback.

4.3. Energy and Transportation

Participant: Nicolas Gast.

This work is mainly done within the Quanticol European project.

Smart urban transport systems and smart grids are two examples of collective adaptive systems. They consist of a large number of heterogeneous entities with decentralised control and varying degrees of complex autonomous behaviour. Within the QUANTICOL project, we develop an analysis tools to help to reason about such systems. Our work relies on tools from fluid and mean-field approximation to build decentralized algorithms that solve complex optimization problems. We focus on two problems: decentralized control of electric grids and capacity planning in vehicle-sharing systems to improve load balancing.

<p style="text-align:center;color:red;">**ROMA Project-Team**</p>

4. Application Domains

4.1. Applications of sparse direct solvers

Sparse direct (e.g., multifrontal solvers that we develop) solvers have a wide range of applications as they are used at the heart of many numerical methods in computational science: whether a model uses finite elements or finite differences, or requires the optimization of a complex linear or nonlinear function, one often ends up solving a system of linear equations involving sparse matrices. There are therefore a number of application fields, among which some of the ones cited by the users of our sparse direct solver MUMPS are: structural mechanics, seismic modeling, biomechanics, medical image processing, tomography, geophysics, electromagnetism, fluid dynamics, econometric models, oil reservoir simulation, magneto-hydro-dynamics, chemistry, acoustics, glaciology, astrophysics, circuit simulation, and work on hybrid direct-iterative methods.

STORM Project-Team

4. Application Domains

4.1. Application Domains

The application of our work concerns linear algebra, solvers and fast-multipole methods, in collaboration with other Inria teams and with industry. This allows a wide range of scientific and industrial applications possibly interested in the techniques we propose, in the domain of high performance computing but also in order to compute intensive embedded applications. In terms of direct application, the software developed in the team are used in applications in various fields, ranging from seismic, mechanic of fluids, molecular dynamics, high energy physics or material simulations. Similarly, the domains of image processing and signal processing can take advantage of the expertise and software of the team.

<h2 style="text-align:center; color:red">TADaaM Project-Team</h2>

4. Application Domains

4.1. Mesh-based applications

TADAAM targets scientific simulation applications on large-scale systems, as these applications present huge challenges in terms of performance, locality, scalability, parallelism and data management. Many of these HPC applications use meshes as the basic model for their computation. For instance, PDE-based simulations using finite differences, finite volumes, or finite elements methods operate on meshes that describe the geometry and the physical properties of the simulated objects. This is the case for at least two thirds of the applications selected in the 9th PRACE. call [0], which concern quantum mechanics, fluid mechanics, climate, material physic, electromagnetism, etc.

Mesh-based applications not only represent the majority of HPC applications running on existing supercomputing systems, yet also feature properties that should be taken into account to achieve scalability and performance on future large-scale systems. These properties are the following:

Size   Datasets are large: some meshes comprise hundreds of millions of elements, or even billions.

Dynamicity   In many simulations, meshes are refined or coarsened at each time step, so as to account for the evolution of the physical simulation (moving parts, shockwaves, structural changes in the model resulting from collisions between mesh parts, etc.).

Structure   Many meshes are unstructured, and require advanced data structures so as to manage irregularity in data storage.

Topology   Due to their rooting in the physical world, meshes exhibit interesting topological properties (low dimensionality embedding, small maximum degree, large diameter, etc.). It is very important to take advantage of these properties when laying out mesh data on systems where communication locality matters.

All these features make mesh-based applications a very interesting and challenging use-case for the research we want to carry out in this project. Moreover, we believe that our proposed approach and solutions will contribute to enhance these applications and allow them to achieve the best possible usage of the available resources of future high-end systems.

---

[0]http://www.prace-ri.eu/prace-9th-regular-call/

ASCOLA Project-Team

4. Application Domains

4.1. Enterprise Information Systems and Services

Large IT infrastructures typically evolve by adding new third-party or internally-developed components, but also frequently by integrating already existing information systems. Integration frequently requires the addition of glue code that mediates between different software components and infrastructures but may also consist in more invasive modifications to implementations, in particular to implement crosscutting functionalities. In more abstract terms, enterprise information systems are subject to structuring problems involving horizontal composition (composition of top-level functionalities) as well as vertical composition (reuse and sharing of implementations among several top-level functionalities). Moreover, information systems have to be more and more dynamic.

Service-Oriented Computing (SOC) that is frequently used for solving some of the integration problems discussed above. Indeed, service-oriented computing has two main advantages:

- Loose-coupling: services are autonomous: they do not require other services to be executed;
- Ease of integration: Services communicate over standard protocols.

Our current work is based on the following observation: similar to other compositional structuring mechanisms, SOAs are subject to the problem of crosscutting functionalities, that is, functionalities that are scattered and tangled over large parts of the architecture and the underlying implementation. Security functionalities, such as access control and monitoring for intrusion detection, are a prime example of such a functionality in that it is not possible to modularize security issues in a well-separated module. Aspect-Oriented Software Development is precisely an application-structuring method that addresses in a systemic way the problem of the lack of modularization facilities for crosscutting functionalities.

We are considering solutions to secure SOAs by providing an aspect-oriented structuring and programming model that allows security functionalities to be modularized. Two levels of research have been identified:

- Service level: as services can be composed to build processes, aspect weaving will deal with the orchestration and the choreography of services.
- Implementation level: as services are abstractly specified, aspect weaving will require to extend service interfaces in order to describe the effects of the executed services on the sensitive resources they control.

4.2. Capacity Planning in Cloud, Fog and Edge Computing

Cloud and more recently Fog and Edge computing platforms aim at delivering large capacities of computing power. These capacities can be used to improve performance (for scientific applications) or availability (e.g., for Internet services hosted by datacenters). These distributed infrastructures consist of a group of coupled computers that work together and may be spread across a LAN (cluster), across a the Internet (Fog/Edge). Due to their large scale, these architectures require permanent adaptation, from the application to the system level and call for automation of the corresponding adaptation processes. We focus on self-configuration and self-optimization functionalities across the whole software stack: from the lower levels (systems mechanisms such as distributed file systems for instance) to the higher ones (i.e. the applications themselves such as clustered servers or scientific applications).

In 2017, we have been consolidating our expertise around the OpenStack ecosystem. We proposed in particular EnOS, a dedicated framework to conduct performance analyses of OpenStack at large-scale ina reproducible manner. The framework enables researchers to conduct experiments in an automize manner on top of different testbeds such as Grid'5000 and Chameleon. see Sec. 7.1 .

In the energy field, we have designed a set of techniques, named Optiplace, for cloud management with flexible power models through constraint programming. OptiPlace supports external models, named views. Specifically, we have developed a power view, based on generic server models, to define and reduce the power consumption of a datacenter's physical servers. We have shown that OptiPlace behaves at least as good as our previous system, Entropy, requiring as low as half the time to find a solution for the constrained-based placement of tasks for large datacenters.

4.3. Pervasive Systems

Pervasive systems are another class of systems raising interesting challenges in terms of software structuring. Such systems are highly concurrent and distributed. Moreover, they assume a high-level of mobility and context-aware interactions between numerous and heterogeneous devices (laptops, PDAs, smartphones, cameras, electronic appliances...). Programming such systems requires proper support for handling various interfering concerns like software customization and evolution, security, privacy, context-awareness... Additionally, service composition occurs spontaneously at runtime.

Like Pervasive systems, Internet of thing is a major theme of these last ten years. Many research works has been led on the whole chain, from communicating sensors to big data management, through communication middlewares.

The more a sensor networks senses various data, the more the users panel is heterogeneous. Such an heterogeneity leads to a major problem about data modeling: for each user, to aim at precisely addressing his needs and his needs only; ie to avoid a data representation which would overwhelm the user with all the data sensed from the network, regardless if he needs it or not. To leverage this issue, we propose a multitree modeling for sensor networks which addresses each of these specific usages.With this modeling comes a domain specific language (DSL) which allows users to manipulate, parse and aggregate information from the sensors.

DIVERSE Project-Team

4. Application Domains

4.1. From Embedded Systems to Service Oriented Architectures

From small embedded systems such as home automation products or automotive systems to medium sized systems such as medical equipment, office equipment, household appliances, smart phones; up to large Service Oriented Architectures (SOA), building a new application from scratch is no longer possible. Such applications reside in (group of) machines that are expected to run continuously for years without unrecoverable errors. Special care has then to be taken to design and validate embedded software, making the appropriate trade-off between various extra-functional properties such as reliability, timeliness, safety and security but also development and production cost, including resource usage of processor, memory, bandwidth, power, etc.

Leveraging ongoing advances in hardware, embedded software is playing an evermore crucial role in our society, bound to increase even more when embedded systems get interconnected to deliver ubiquitous SOA. For this reason, embedded software has been growing in size and complexity at an exponential rate for the past 20 years, pleading for a component based approach to embedded software development. There is a real need for flexible solutions allowing to deal at the same time with a wide range of needs (product lines modeling and methodologies for managing them), while preserving quality and reducing the time to market (such as derivation and validation tools).

We believe that building flexible, reliable and efficient embedded software will be achieved by reducing the gap between executable programs, their models, and the platform on which they execute, and by developing new composition mechanisms as well as transformation techniques with a sound formal basis for mapping between the different levels.

Reliability is an essential requirement in a context where a huge number of softwares (and sometimes several versions of the same program) may coexist in a large system. On one hand, software should be able to evolve very fast, as new features or services are frequently added to existing ones, but on the other hand, the occurrence of a fault in a system can be very costly, and time consuming. While we think that formal methods may help solving this kind of problems, we develop approaches where they are kept "behind the scene" in a global process taking into account constraints and objectives coming from user requirements.

Software testing is another aspect of reliable development. Testing activities mostly consist in trying to exhibit cases where a system implementation does not conform to its specifications. Whatever the efforts spent for development, this phase is of real importance to raise the confidence level in the fact that a system behaves properly in a complex environment. We also put a particular emphasis on on-line approaches, in which test and observation are dynamically computed during execution.

<p align="center" style="color:red">FOCUS Project-Team</p>

4. Application Domains

4.1. Ubiquitous Systems

The main application domain for Focus are ubiquitous systems, broadly systems whose distinctive features are: mobility, high dynamicity, heterogeneity, variable availability (the availability of services offered by the constituent parts of a system may fluctuate, and similarly the guarantees offered by single components may not be the same all the time), open-endedness, complexity (the systems are made of a large number of components, with sophisticated architectural structures). In Focus we are particularly interested in the following aspects.

- *Linguistic primitives* for programming dialogues among components.
- *Contracts* expressing the functionalities offered by components.
- *Adaptability and evolvability* of the behaviour of components.
- *Verification* of properties of component systems.
- Bounds on component *resource consumption* (e.g., time and space consumed).

4.2. Service Oriented Computing and Cloud Computing

Today the component-based methodology often refers to Service Oriented Computing. This is a specialized form of component-based approach. According to W3C, a service-oriented architecture is "a set of components which can be invoked, and whose interface descriptions can be published and discovered". In the early days of Service Oriented Computing, the term services was strictly related to that of Web Services. Nowadays, it has a much broader meaning as exemplified by the XaaS (everything as a service) paradigm: for example, based on modern virtualization technologies, cloud computing offers the possibility to build sophisticated service systems on virtualized infrastructures accessible from everywhere and from any kind of computing device. Such infrastructures are usually examples of sophisticated service oriented architectures that, differently from traditional service systems, should also be capable to elastically adapt on demand to the user requests.

INDES Project-Team (section vide)

PHOENIX Project-Team

4. Application Domains

4.1. Internet of Things

Participants: Charles Consel, Nic Volanschi.

The Internet of Things (IoT) has become a reality with the emergence of Smart Cities, populated with large amounts of smart objects which are used to deliver a range of citizen services (e.g., security, well being, etc.) The IoT paradigm relies on the pervasive presence of smart objects or "things", which raises a number of new challenges in the software engineering domain.

We introduce a design-driven development approach that is dedicated to the domain of orchestration of masses of sensors. The developer declares what an application does using a domain-specific language (DSL), named DiaSwarm. Our compiler processes domain-specific declarations to generate a customized programming framework that guides and supports the programming phase.

DiaSwarm addresses the main phases of an application orchestrating masses of sensors.

Service discovery   Standard service discovery at the individual object level does not address the needs of applications orchestrating large numbers of smart objects. Instead, a high-level approach which provides constructs to specifying subsets of interest is needed. Our approach allows developers to introduce application-specific concepts (e.g., regrouping parking spaces into lots or districts) at the design time and then these can be used to express discovery operations. Following our design-driven development approach, these concepts are used to generate code to support and guide the programming phase.

Data gathering   Applications need to acquire data from a large number of objects through a variety of delivery models. For instance, air pollution sensors across a city may only push data to the relevant applications when pollution levels exceed tolerated levels. Tracking sensors, however, might determine the location of vehicles and send the acquired measurements to applications periodically (e.g., 10 min. intervals). Data delivery models need to be introduced at design time since they have a direct impact on the application's program structure. In doing so, the delivery models used by an application can be checked against sensor features early in the development process.

Data processing   Data that is generated from hundreds of thousands of objects and accumulated over a period of time calls for efficient processing strategies to ensure the required performance is attained. Our approach allows for an efficient implementation of the data processing stage by providing the developer with a framework based on the MapReduce [34] programming model which is intended for the processing of large data sets.

4.2. Assistive computing in the home

Participants: Quentin Barlas, Antoine Riche, Adrien Carteron, Rafik Belloum, Audrey Landuran, Julien Durand, Charlotte Froger, Stéphanie Giraud, Pauline Fontagne, Flora Gallet, Amandine Desrozier, Ludovic Fornasari, Léo Mendiboure, Hélène Sauzéon, Charles Consel, Bernard N'Kaoua, Bernard Serpette, Nic Volanschi.

In this avenue of research, we have been developing a systemic approach to introducing an assisted living platform for the home of older adults. To do so, we formed an interdisciplinary team that allows (1) to identify the user needs from a gerontological and psychological viewpoint; (2) to propose assistive applications designed by human factors and HCI experts, in collaboration with caregivers and users; (3) to develop and test applications designed and developed by software engineers; (4) to conduct a field study to assess the benefits of the platform and assistive applications, in collaboration with caregivers, by deploying the system at the actual homes.

Our research activities for assistive computing in the home are conducted under the *HomeAssist* project. This work takes the form of a platform offering an online catalog of assistive applications that orchestrate an open-ended set of networked objects. Our platform leverages DiaSuite to quickly and safely develop applications at a high level.

Our scientific achievements include the design principles of our platform, its key features to effectively assist individuals in their home, field studies to validate HomeAssist, the expansion of HomeAssist to serve individuals with ID, and the technology transfer of HomeAssist. Note that a complete presentation of this work, from a Cognitive Science perspective, is given in the doctoral thesis of Lucile Dupuy published last year (2016).

4.2.1. *Project-team positioning*

There is a range of platforms for assisted living aimed at older adults that have been developed for more than a decade. Most of these platforms are used in a setting where participants come to a research apartment to perform certain tasks. This setting makes it difficult to assess user acceptance and satisfaction of the proposed approaches because the user does not interact with the technology on a daily basis, over a period of time. Furthermore, older adults adopt routines to optimize their daily functioning at home. This situation calls for field studies in a naturalistic setting to strengthen the evaluation of assisted living platforms.

HomeAssist innovates in that it supports independent living across the activities of daily living and is validated by field studies in naturalistic setting.

4.3. Assistive computing on-the-go

Participants: Cécile Mazon, Léo Mendiboure, Hélène Sauzéon, Charles Consel.

We conduct research on assistive computing supported by mobile devices such as smart phones and tablets. Both research projects presented in this section are supported by tablets and leverage their functionalities to guide users with cognitive challenges performing activities and tasks, whether in mainstream schools to support inclusion or in residential settings to support their autonomy. The mobile nature of tablets allows to envision such devices as supporting users with cognitive challenges across a range of environments.

Many research projects bring cognitive-support applications to users based on tablets and smartphones. However, few projects equip users with such devices in actual mainstream environments, including stakeholders in the design process and targeting an autonomous usage of assistive applications. An additional originality of our approach is our interdisciplinary approach that allows us to integrate key psychological dimensions in our design, such as self-determination.

4.4. Life Plan

Participants: Audrey Landuran, Gregory Lecouvey, Léo Mendiboure, Quentin Barlas, Bernard N'Kaoua.

Elaboration of life plan is a major stage in the developmental trajectory of people with intellectual deficiency and is based on several capabilities : make choices, express them, be aware of their consequences, etc. However, these various capabilities can raise problems for people with Down syndrome. The aim of this project has been to design and validate a digital assistant that allows to help the individuals to make choices, to plan for the future and to define their life plan. The user centered methodology employed has allowed to involve the people with Down syndrome, their families and their caregivers in all stages of design. Different validations achieved attest for the accessible, friendly, funny character of the assistant and its ability to promote the expression of the life plan in accordance with individuals' choices, wishes and desires.

The digital assistant is now available online at : http://www.monprojetdevie.trisomie21-france.org/.

Figure 2. Sample screenshots from the Life Plan website.

<p align="center">**RMOD Project-Team**</p>

4. Application Domains

4.1. Programming Languages and Tools

Many of the results of RMoD are improving programming languages or development tools for such languages. As such the application domain of these results is as varied as the use of programming languages in general. Pharo, the language that RMoD develops, is used for a very broad range of applications. From pure research experiments to real world industrial use (the Pharo Consortium, http://consortium.pharo.org, has more than 25 company members).

Examples are web applications, server backends for mobile applications or even graphical tools and embedded applications

4.2. Software Reengineering

Moose is a language-independent environment for reverse and re-engineering complex software systems. Moose provides a set of services including a common meta-model, metrics evaluation and visualization. As such Moose is used for analyzing software systems to support understanding and continuous development as well as software quality analysis.

TACOMA Team

4. Application Domains

4.1. Pervasive applications in Smart Building

A Smart Building is a living space equipped with information-and-communication-technology (ICT) devices conceived to collaborate in order to anticipate and respond to the needs of the occupants, working to promote their comfort, convenience, security and entertainment while preserving their natural interaction with the environment.

The idea of using the Pervasive Computing paradigm in the Smart Building domain is not new. However, the state-of-the-art solutions only partially adhere to its principles. Often the adopted approach consists in a heavy deployment of sensor nodes, which continuously send a lot of data to a central elaboration unit, in charge of the difficult task of extrapolating meaningful information using complex techniques. This is a *logical approach*. TACOMA proposed instead the adoption of a *physical approach*, in which the information is spread in the environment, carried by the entities themselves, and the elaboration is directly executed by these entities "inside" the physical space. This allows performing meaningful exchanges of data that will thereafter need a less complicated processing compared to the current solutions. The result is a smart environment that can, in an easier and better way, integrate the context in its functioning and thus seamlessly deliver more useful and effective user services. Our contribution aims at implementing the physical approach in a smarter environment, showing a solution for improving both comfort and energy savings.

4.2. Metamorphic House

The motivation for metamorphic houses is that many countries, including France, are going through socio-demographic evolutions, like growth of life expectancy and consequent increase in the number of elderly people, urbanization and resource scarcity. Households experience financial restrictions, while housing costs increase with the raise of real estate and energy prices [6].

Important questions arise concerning the future of housing policies and ways of living. We observe novel initiatives like participative housing and developing behaviors, including house-sharing, teleworking and longer stay of children in parents' homes.

To tackle the challenges raised by these emerging phenomena, future homes will have to be modular, upgradeable, comfortable, sparing of resources. They should be integrated in the urban context and exchange information with other homes, contribute to reducing the distances to be covered daily and respect the characteristics of the territory where they are located.

To reach these goals, metamorphic domestic environments will modify their shape and behavior to support activities and changes in life cycle of occupants, increase comfort and optimize the use of resources. Thanks to Information and Communication Technologies (ICT) and adaptive building elements, the same physical spaces will be transformed for different uses, giving inhabitants the illusion of living in bigger, more adapted and more comfortable places.

4.3. Automation in Smart City

The domain of Smart Cities is still young but it is already a huge market which attract number of companies and researchers. It is also multi-fold as the words "smart city" gather multiple meanings. Among them one of the main responsibilities of a city, is to organize the transportation of goods and people. In intelligent transportation systems (ITS), ICT technologies have been involved to improve planification and more generally efficiency of journeys within the city. We are interested in the next step where efficiency would be improved locally relying on local interactions between vehicles, infrastructure and people (smartphones).

For the future autonomous vehicle are now in the spotlight, since a lot of works has been done in recent years in automotive industry as well as in academic research centers. Such unmanned vehicle could strongly impact the organisation of the transportation in our cities. However, due to the lack of a definition of what is an "autonomous" vehicle it remains still difficult to see how these vehicles will interact with their environment (eg. road, smart city, houses, grid, etc.). From augmented perception to fully cooperative automated vehicle, the autonomy covers various realities in terms of interaction the vehicle relies on. The extended perception relies on communication between the vehicle and surrounding roadside equipments. This help the driving system to build and maintain an accurate view of the environment. But at this first stage the vehicle only uses its own perception to make its decisions. At a second stage, it will take advantages of local interaction with other vehicles through car-to-car communications to elaborate a better view of its environment. Such "cooperative autonomy" does not try to reproduce the human behavior anymore, it strongly relies on communication between vehicles and/or with the infrastructure to make decision and to acquire information on the environment. Part of the decision could be centralized (almost everything for an automatic metro) or coordinated by a roadside component. The decision making could even be fully distributed but this put high constraints on the communications. Automated vehicles are just an exemple of smart city automated processes that will have to share information within the surrounding to make their decisions.

4.4. Pervasive applications in uncontrolled environnements

Some limitations of existing RFID technology become challenging: unlike standard RFID application scenarios, pervasive computing often involves uncontrolled environment for RFID, where tags and reader have to operate in much more difficult situations that those usually encountered or expected for classical RFID systems.

RFID technology is to avoid missing tags when reading multiple objects, as reading reliability is affected by various effects such shadowing or wave power absorption by some materials. The usual applications of RFID operate in a controlled environment in order to reduce the risk of missing tags while scanning objects.

In pervasive computing applications, a controlled reading environment is extremely difficult to achieve, as one of the principle is to enhance existing processes "in situ", unlike the controlled conditions that can be found in industrial processes. Consider for example a logistic application, where RFID tags could be used on items inside a package in order to check for its integrity along the shipping process. Tags would likely be placed randomly on items inside the package, and reading conditions would be variable depending on where the package is checked.

RFID operation in uncontrolled environments is challenging because RFID performance is affected by multiple parameters, in particular:

- Objects materials (on which tags are attached to),
- Materials in the surrounding environment,
- RFID frequency spectrum,
- Antenna nature and placement with respect to the tags.

In controlled environment, the difficulty to read tags can be limited by using the appropriate parameters to maximize the RFID performance for the application. But in many cases, it is needed to read large number of objects of various nature, arranged randomly in a given area or container. **Most pervasive computing applications fall in this context**.

AGORA Team

4. Application Domains

4.1. Smart Cities

One major characteristic of modern societies is that they are prevalently urban. In coherence, the contributions of the Agora team are in particular applied to provide solutions tailored to the emergence of the Internet of Things (IoT) and to Smart Cities applications. A major motivation of the team is the forthcoming explosion of the number of connected devices. In particular, low cost - small data devices are supposed to be densely deployed in our environment, fostering the interest for a convergence of the traditional wireless networking paradigms.

Smart City is a constantly reshaped concept, embracing the future of dense metropolitan areas, with references to efficient and sustainable infrastructure, improving citizens' quality of life and protecting the environment. A consensus on the Smart City philosophy is however that it will be primarily achieved by leveraging a clever integration of Information and Communication Technologies (ICT) in the urban tissue. Indeed, ICTs are enabling an evolution from the current duality between the real world and its digitized counterpart to a continuum in which digital contents and applications are seamlessly interacting with classical infrastructures and services. Smart Cities are often described by the digital services that should be provided which are inherently dependent on dense measurements of the city environment and activities, the collection of these data, their processing into information, and their redistribution. The networking infrastructure plays therefore a critical role in enabling advanced services, in particular the wireless infrastructure supporting density and mobility.

From a wireless networking viewpoint, the digitization of cities can be seen as a paradigm shift extending the Internet of Things (IoT) to a citizen-centric model in order to leverage the massive data collected by pervasive sensors, connected mobiles or fixed devices, and social applications.

<h2 style="text-align:center;color:red;">COATI Project-Team</h2>

4. Application Domains

4.1. Telecommunication Networks

COATI is mostly interested in telecommunications networks. We focus on the design and management of heterogeneous physical and logical networks. The project has kept working on the design of backbone networks (optical networks, radio networks, IP networks). We also study routing algorithms such as dynamic and compact routing schemes, as we did in the context of the FP7 EULER led by Alcatel-Lucent Bell-Labs (Belgium), and the evolution of the routing in case of any kind of topological modifications (maintenance operations, failures, capacity variations, etc.). However, the fields of Software Defined Networks and Network Function Virtualization are growing in importance in our studies.

4.2. Other Domains

Our combinatorial tools may be well applied to solve many other problems in various areas (transport, biology, resource allocation, chemistry, smart-grids, speleology, etc.) and we intend to collaborate with experts of these other domains.

For instance, we collaborate with EP ABS (Algorithms Biology Structure) from Sophia Antipolis on problems from Structural Biology (see Section 7.2.1.5). In the area of transportation networks, we have started a collaboration with SME Instant-System on dynamic car-pooling combined with multi-modal transportation systems. This collaboration will be strengthen in the near future with the support of an ANR project (starting January 2018).

DANTE Project-Team

4. Application Domains

4.1. Life Science & Health

In parallel to the advances in modern medicine, health sciences and public health policy, epidemic models aided by computer simulations and information technologies offer an increasingly important tool for the understanding of transmission dynamics and of epidemic patterns. The increased computational power and use of Information and Communication Technologies make feasible sophisticated modelling approaches augmented by detailed in vivo data sets, and allow to study a variety of possible scenarios and control strategies, helping and supporting the decision process at the scientific, medical and public health level. The research conducted in the DANTE project finds direct applications in the domain of LSH since modelling approaches crucially depend on our ability to describe the interactions of individuals in the population. In the MOSAR/iBird project we are collaborating with the team of Pr. Didier Guillemot (Inserm/Institut. Pasteur/Université de Versailles). Within the TUBEXPO and ARIBO projects, we are collaborating with Pr. Jean-Christopge Lucet (Professeur des université Paris VII, Praticien hospitalier APHP).

4.2. Network Science / Complex networks

In the last ten years the science of complex networks has been assigned an increasingly relevant role in defining a conceptual framework for the analysis of complex systems. Network science is concerned with graphs that map entities and their interactions to nodes and links. For a long time, this mathematical abstraction has contributed to the understanding of real-world systems in physics, computer science, biology, chemistry, social sciences, and economics. Recently, however, enormous amounts of detailed data, electronically collected and meticulously catalogued, have finally become available for scientific analysis and study. This has led to the discovery that most networks describing real world systems show the presence of complex properties and heterogeneities, which cannot be neglected in their topological and dynamical description. This has called forth a major effort in developing the methodology to characterise the topology and temporal behaviour of complex networks, to describe the observed structural and temporal heterogeneities, to detect and measure emerging community structure, to see how the functionality of networks determines their evolving structure, and to determine what kinds of correlations play a role in their dynamics. All these efforts have brought us to a point where the science of complex networks has become advanced enough to help us to disclose the deeper roles of complexity and gain understanding about the behaviour of very complicated systems.

In this endeavour the DANTE project targets the study of dynamically evolving networks, concentrating on questions about the evolving structure and dynamical processes taking place on them. During the last year we developed developed several projects along these lines concerning three major datasets:

- Mobile telephony data: In projects with academic partners and Grandata we performed projects based on two large independent datasets collecting the telephone call and SMS event records for million of anonymised individuals. The datasets record the time and duration of mobile phone interactions and some coarse grained location and demographic data for some users. In addition one of the dataset is coupled with anonymised bank credit information allowing us to study directly the socioeconomic structure of a society and how it determines the communication dynamics and structure of individuals.

- Skype data: Together with Skype Labs/STACC and other academic groups we were leading projects in the subject of social spreading phenomena. These projects were based on observations taken from a temporally detailed description of the evolving social network of (anonymised) Skype users registered between 2003 and 2011. This data contains dates of registration and link creation together with gradual information about their location and service usage dynamics.

- Twitter data: In collaboration with ICAR-ENS Lyon we collected a large dataset about the microblogs and communications of millions of Twitter users in the French Twitter space. This data allows us to follow the spreading of fads/opinions/hashtags/ideas and more importantly linguistic features in online communities. The aim of this collaboration is to set the ground for a quantitative framework studying the evolution of linguistic features and dialects in an social-communication space mediated by online social interactions.

DIANA Project-Team (section vide)

<p style="text-align:center">**DIONYSOS Project-Team**</p>

4. Application Domains

4.1. Networking

Our global research effort concerns networking problems, both from the analysis point of view, and around network design issues. Specifically, this means the IP technology in general, with focus on specific types of networks seen at different levels: wireless systems, optical infrastructures, peer-to-peer architectures, Software Defined Networks, Content Delivery Networks, Content-Centric Networks, clouds.

A specific aspect of network applications and/or services based on video or voice content, is our PSQA technology, able to measure the Perceptual Quality automatically and in real time. PSQA provides a MOS value as close as it makes sense to the value obtained from subjective testing sessions. The technology has been tested in many environments, including one way communications as, for instance, in video streaming, and bi-directional communications as in IP telephony, UDP- or TCP-based systems, etc. It has already served in many collaborative projects as the measuring tool used.

4.2. Stochastic modeling

Many of the techniques developed at Dionysos are related to the analysis of complex systems in general, not only in telecommunications. For instance, our Monte Carlo methods for analyzing rare events have been used by different industrial partners, some of them in networking but recently also by companies building transportation systems. We develop methods in different areas: numerical analysis of stochastic models, bound computations in the same area, Discrete Event Simulation, or, as just mentioned, rare event analysis.

DYOGENE Project-Team

4. Application Domains

4.1. Physical communication networks

Internet, wireless, mobile, cellular networks.

4.2. Abstract networks

Social interactions, human communities, economic networks.

4.3. Power grids

Energy networks.

EVA Project-Team

4. Application Domains

4.1. Industrial Process Automation

Wireless networks have become ubiquitous and are an integral part of our daily lives. These networks are present in many application domains; the most important are detailed in this section.

Networks in industrial process automation typically perform **monitoring and control** tasks. Wired industrial communication networks, such as HART [0], have been around for decades and, being wired, are highly reliable. Network administrators tempted to "go wireless" expect the same reliability. Reliable process automation networks – especially when used for control – often impose stringent latency requirements. Deterministic wireless networks can be used in critical systems such as control loops, however, the unreliable nature of the wireless medium, coupled with their large scale and "ad-hoc" nature raise some of the most important challenges for low-power wireless research over the next 5-10 years.

Through the involvement of team members in standardization activities, the protocols and techniques will be proposed for the standardization process with a view to becoming the *de-facto* standard for wireless industrial process automation. Besides producing top level research publications and standardization activities, EVA intends this activity to foster further collaborations with industrial partners.

4.2. Environmental Monitoring

Today, outdoor WSNs are used to monitor vast rural or semi-rural areas and may be used to detect fires. Another example is detecting fires in outdoor fuel depots, where the delivery of alarm messages to a monitoring station in an upper-bounded time is of prime importance. Other applications consist in monitoring the snow melting process in mountains, tracking the quality of water in cities, registering the height of water in pipes to foresee flooding, etc. These applications lead to a vast number of technical issues: deployment strategies to ensure suitable coverage and good network connectivity, energy efficiency, reliability and latency, etc.

We work on such applications in an associate team "REALMS" comprising members from EVA, the university of Berkeley and the university of Michigan.

4.3. The Internet of Things

The general agreement is that the Internet of Things (IoT) is composed of small, often battery-powered objects which measure and interact with the physical world, and encompasses smart home applications, wearables, smart city and smart plant applications.

It is absolutely essential to (1) clearly understand the limits and capabilities of the IoT, and (2) develop technologies which enable user expectation to be met.

The EVA team is dedicated to understanding and contributing to the IoT. In particular, the team will maintain a good understanding of the different technologies at play (Bluetooth, IEEE 802.15.4, WiFi, cellular), and their trade-offs. Through scientific publications and other contributions, EVA will help establishing which technology best fits which application.

---

[0]Highway Addressable Remote Transducer

4.4. Military, Energy and Aerospace

Through the HIPERCOM project, EVA has developed cutting-edge expertise in using wireless networks for military, energy and aerospace applications. Wireless networks are a key enabling technology in the application domains, as they allow physical processes to be instrumented (e.g. the structural health of an airplane) at a granularity not achievable by its wired counterpart. Using wireless technology in these domains does however raise many technical challenges, including end-to-end latency, energy-efficiency, reliability and Quality of Service (QoS). Mobility is often an additional constraint in energy and military applications. Achieving scalability is of paramount importance for tactical military networks, and, albeit to a lesser degree, for power plants. EVA will work in this domain.

Smart cities share the constraint of mobility (both pedestrian and vehicular) with tactical military networks. Vehicular Ad-hoc NETworks (VANETs) will play an important role in the development of smarter cities.

The coexistence of different networks operating in the same radio spectrum can cause interference that should be avoided. Cognitive radio provides secondary users with the frequency channels that are temporarily unused (or unassigned) by primary users. Such opportunistic behavior can also be applied to urban wireless sensor networks. Smart cities raise the problem of transmitting, gathering, processing and storing big data. Another issue is to provide the right information at the place where it is most needed.

4.5. Emergency Applications

In an "emergency" application, heterogeneous nodes of a wireless network cooperate to recover from a disruptive event in a timely fashion, thereby possibly saving human lives. These wireless networks can be rapidly deployed and are useful to assess damage and take initial decisions. Their primary goal is to maintain connectivity with the humans or mobile robots (possibly in a hostile environment) in charge of network deployment. The deployment should ensure the coverage of particular points or areas of interest. The wireless network has to cope with pedestrian mobility and robot/vehicle mobility. The environment, initially unknown, is progressively discovered and may contain numerous obstacles that should be avoided. The nodes of the wireless network are usually battery-powered. Since they are placed by a robot or a human, their weight is very limited. The protocols supported by these nodes should be energy-efficient to maximize network lifetime. In such a challenging environment, sensor nodes should be replaced before their batteries are depleted. It is therefore important to be able to accurately determine the battery lifetime of these nodes, enabling predictive maintenance.

4.6. Types of Wireless Networks

The EVA team will distinguish between opportunistic communication (which takes advantage of a favorable state) and = collaborative communication (several entities collaborate to reach a common objective). Furthermore, determinism can be required to schedule medium access and node activity, and to predict energy consumption.

In the EVA project, we will propose **self-adaptive wireless networks** whose evolution is based on:

- optimization to minimize a single or multiple objective functions under some constraints (e.g. interference, or energy consumption in the routing process).
- machine learning to be able to predict a future state based on past states (e.g. link quality in a wireless sensor network) and to identify tendencies.

The types of wireless networks encountered in the application domains can be classified in the following categories.

4.6.1. Wireless Sensor and Mesh Networks

Standardization activities at the IETF have defined an "upper stack" allowing low-power mesh networks to be seamlessly integrated in the Internet (6LoWPAN), form multi-hop topologies (RPL), and interact with other devices like regular web servers (CoAP).

Major research challenges in sensor networks are mostly related to (predictable) power conservation and efficient multi-hop routing. Applications such as monitoring of mobile targets, and the generalization of smart phone devices and wearables, have introduced the need for WSN communication protocols to cope with node mobility and intermittent connectivity.

Extending WSN technology to new application spaces (e.g. security, sports, hostile environments) could also assist communication by seamless exchanges of information between individuals, between individuals and machines, or between machines, leading to the Internet of Things.

4.6.2. Deterministic Low-Power Networks

Wired sensor networks have been used for decades to automate production processes in industrial applications, through standards such as HART. Because of the unreliable nature of the wireless medium, a wireless version of such industrial networks was long considered infeasible.

In 2012, the publication of the IEEE 802.15.4e standard triggered a revolutionary trend in low-power mesh networking: merging the performance of industrial networks, with the ease-of-integration of IP-enabled networks. This integration process is spearheaded by the IETF 6TiSCH working group, created in 2013. A 6TiSCH network implements the IEEE 802.15.4e TSCH protocol, as well as IETF standards such as 6LoWPAN, RPL and CoAP. A 6TiSCH network is synchronized, and a communication schedule orchestrates all communication in the network. Deployments of pre-6TiSCH networks have shown that they can achieve over 99.999% end-to-end reliability, and a decade of battery lifetime.

The communication schedule of a 6TiSCH network can be built and maintained using a centralized, distributed, or hybrid scheduling approach. While the mechanisms for managing that schedule are being standardized by the IETF, which scheduling approach to use, and the associated limits in terms of reliability, throughput and power consumption remains entirely open research questions. Contributing to answering these questions is an important research direction for the EVA team.

4.6.3. MANETs and VANETs

In contrast to routing, other domains in MANETs such as medium access, multi-carrier transmission, quality of service, and quality of experience have received less attention. The establishment of research contracts for EVA in the field of MANETs is expected to remain substantial. MANETs will remain a key application domain for EVA with users such as the military, firefighters, emergency services and NGOs.

Vehicular Ad hoc Networks (VANETs) are arguably one of the most promising applications for MANETs. These networks primarily aim at improving road safety. Radio spectrum has been ring-fenced for VANETs worldwide, especially for safety applications. International standardization bodies are working on building efficient standards to govern vehicle-to-vehicle or vehicle-to-infrastructure communication.

4.6.4. Cellular and Device-to-Device Networks

We propose to initially focus this activity on spectrum sensing. For efficient spectrum sensing, the first step is to discover the links (subcarriers) on which nodes may initiate communications.= In Device-to-Device (D2D) networks, one difficulty is scalability.

For link sensing, we will study and design new random access schemes for D2D networks, starting from active signaling. This will assume the availability of a control channel devoted to D2D neighbor discovery. It is therefore naturally coupled with cognitive radio algorithms (allocating such resources): coordination of link discovery through eNode-B information exchanges can yield further spectrum usage optimization.

<p style="text-align: center; color: red;">FUN Project-Team</p>

4. Application Domains

4.1. Application Domains

The set of applications enabled through FUN and IoT is very large and can apply in every application area. We can thus not be exhaustive but among the most spread applications, we can name every area, event or animal monitoring, understanding and protection. To illustrate this, we may refer to the use cases addressed by our PREDNET project which goals is to equip rhinoceros with smart communicating devices to fight against poaching.

Other field of application is exploration of hostile and/or unknown environment by a fleet of self-organizing robots that cooperate with RFID and sensors to ensure a continue monitoring afterwards.

Also, IoT and FUN ca play a key role in logistics and traceability by relying on the use of sensors or RFID technologies as implemented in our TRACAVERRE project or our collaboration with the start up TRAXENS.

Finally, IoT and FUN leverage a lot of applications in Smart City concept , ranging from parking aid to a better energy consumption going through air quality monitoring, traffic fluidizing etc. (See our CityLab Inria and VITAL projects).

<p style="text-align: center; color: red;">**GANG Project-Team**</p>

4. Application Domains

4.1. Large scale networks

Application domains include evaluating Internet performances, the design of new peer-to-peer applications, enabling large scale networks, and developping tools for transportation networks.

INFINE Project-Team (section vide)

MADYNES Team

4. Application Domains

4.1. Mobile and constrained networks

The results coming out from MADYNES can be applied to any dynamic infrastructure that contributes to the delivery of value added services. While this is a potentially huge application domain, we mainly focus at the network level on mobile devices and Internet of Things. We are investigating the provisioning, monitoring, configuration and performance management issues.

4.2. Dynamic services infrastructures

At the service level, dynamics is also increasing very fast. We apply the results of our work on autonomous management on infrastructures which support dynamic composition and for which self-instrumentation and management automation is required. The target service environments are:

- sensor networks,
- cyber-physical systems,
- information centric networks,
- distributed cloud environnements,
- smart environments.

<p style="text-align:center">**NEO Project-Team**</p>

4. Application Domains

4.1. Network Science

Network Science is a multidisciplinary body of knowledge, principally concerned with the emergence of global properties in a network of individual agents, from the "local" properties of this network, namely, the way agents interact with each other. The central model of "networks" is the graph (of Graph Theory/Operations Research), with nodes representing the different entities managing information and taking decisions, and the links representing the fact that entities interact, or not. Links are usually equipped with a "weight" that measures the intensity of interaction. Adding evolution rules to this quite elementary representation leads to dynamic network models, the properties of which Network Science tries to analyze.

A classical example of properties sought in networks is the famous "six degrees of separation" (or "small world") property: how and why does it happen so frequently? Another ubiquitous property of real-life networks is the Zipf or "scale-free" distribution for degrees. Some of these properties, when properly exploited, lead to successful business opportunities: just consider the PageRank algorithm of Google, which miraculously connects the relevance of some Web information with the relevance of the other information that points to it.

4.2. Network Engineering

In its primary acceptation, Network Science involves little or no engineering: phenomena are assumed to be "natural" and emerge without intervention. However, the idea comes fast to intervene in order to modify the outcome of the phenomenon. This is where NEO is positioned. Beyond the mostly descriptive approach of Network Science, we aim at using the techniques of Operations Research so as to engineer complex networks.

To quote just two examples: controlling the spread of diseases through a "network" of people is of primarily interest for mankind. Similarly, controlling the spread of information or reputation through a social network is of great interest in the Internet. Precisely: given the impact of web visibility on business income, it is tempting (and quite common) to manipulate the graph of the web by adding links so as to drive the PageRank algorithm to a desired outcome.

Another interesting example is the engineering of community structures. Recently, thousands of papers have been written on the topic of community *detection* problem. In most of the works, the researchers propose methods, most of the time, heuristics, for detecting communities or dense subgraphs inside a large network. Much less effort has been put in the understanding of community formation process and even much less effort has been dedicated to the question of how one can influence the process of community formation, e.g. in order to increase overlap among communities and reverse the fragmentation of the society.

Our ambition for the medium term is to reach an understanding of the behavior of complex networks that will make us capable of influencing or producing a certain property in said network. For this purpose, we will develop families of models to capture the essential structure, dynamics, and uncertainty of complex networks. The "solution" of these models will provide the correspondence between metrics of interest and model parameters, thus opening the way to the synthesis of effective control techniques.

In the process of tackling real, very large size networks, we increasingly deal with large graph data analysis and the development of decision techniques with low algorithmic complexity, apt at providing answers from large datasets in reasonable time.

RAP2 Team (section vide)

SOCRATE Project-Team (section vide)