Imprimir esta página

Voltar ao Projecto

Ocultar todos os formulários
Hide all forms

Referência do projecto
Project reference
PTDC/EIA/76114/2006

---

1. Identificação do projecto
1. Project description

Financiamento solicitado
Requested funding
105.288,00 Euros

Área científica principal
Main Area
Computer Engineering
Área científica Secundária
Secondary area
(vazio)
(void)
Título do projecto (em português)
Project title (in portuguese)
LiveFeeds - Disseminação P2P de Conteúdos Web Syndication
Título do projecto (em inglês)
Project title (in english)
LiveFeeds - P2P Dissemination of Web Syndication Content

Palavra-chave 1
disseminação de conteúdos em larga-escala
Keyword 1
large-scale content dissemination
Palavra-chave 2
algoritmos P2P
Keyword 2
P2P algorithms
Palavra-chave 3
notificação em larga-escala
Keyword 3
large-scale notification
Palavra-chave 4
modelo editor/assinante
Keyword 4
publish/subscribe

Objectivos sócio-económicos
Socio-economic objectives
General Advancement Of Knowledge

Data de início do projecto
Starting date
01-01-2007
Duração do projecto em meses
Duration in months
36

---

2. Instituições participantes
2. Participating institutions

Instituição Proponente
Principal Contractor
Fundação da Faculdade de Ciências e Tecnologia (FFCT/FCT/UNL)
Campus da Caparica
2829-516Caparica

Instituições Participantes
Participating Institutions
(vazio)
(void)

Unidade de Investigação
Principal Research Unit
Centro de Informática e Tecnologias da Informação (CITI/FCT/UNL)

Quinta da Torre
2829-516Monte de Caparica

Instituição de Acolhimento
Host Institution
Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (FCT/UNL)
Campus de Caparica
2829-516Monte de Caparica

3. Orçamento
3. Proposal Budget

Instituição Proponente
Principal Contractor

Fundação da Faculdade de Ciências e Tecnologia

DESCRIÇÃO DESCRIPTION	**2007**	**2008**	**2009**	**2010**	**2011**	**TOTAL**
Recursos Humanos Human resources	18.084,00	23.172,00	18.084,00	0,00	0,00	**59.340,00**
Missões Missions	2.400,00	4.400,00	4.400,00	0,00	0,00	**11.200,00**
Consultores Consultants	0,00	0,00	0,00	0,00	0,00	**0,00**
Aquisição de serviços e manutenção Acquisition of services and maintenance	500,00	750,00	750,00	0,00	0,00	**2.000,00**
Outras despesas correntes Other current expenses	250,00	500,00	500,00	0,00	0,00	**1.250,00**
Despesas gerais Overheads	4.787,00	6.964,00	5.797,00	0,00	0,00	**17.548,00**
TOTAL DESPESAS CORRENTES **TOTAL CURRENT EXPENSES**	**26.021,00**	**35.786,00**	**29.531,00**	**0,00**	**0,00**	**91.338,00**
Equipamento Equipment	2.700,00	6.000,00	5.250,00	0,00	0,00	**13.950,00**
TOTAL	**28.721,00**	**41.786,00**	**34.781,00**	**0,00**	**0,00**	**105.288,00**

Instituições Participantes
Participating Institutions
(Não se encontram registadas Instituições Participantes para este projecto)
(No Participating Institution has been registered for this project)

Orçamento Global
Global budget

DESCRIÇÃO DESCRIPTION	**2007**	**2008**	**2009**	**2010**	**2011**	**TOTAL**
Recursos Humanos Human resources	18.084,00	23.172,00	18.084,00	0,00	0,00	**59.340,00**
Missões Missions	2.400,00	4.400,00	4.400,00	0,00	0,00	**11.200,00**
Consultores Consultants	0,00	0,00	0,00	0,00	0,00	**0,00**
Aquisição de serviços e manutenção Acquisition of services and maintenance	500,00	750,00	750,00	0,00	0,00	**2.000,00**
Outras despesas correntes Other current expenses	250,00	500,00	500,00	0,00	0,00	**1.250,00**
Despesas gerais Overheads	4.787,00	6.964,00	5.797,00	0,00	0,00	**17.548,00**
TOTAL DESPESAS CORRENTES **TOTAL CURRENT EXPENSES**	**26.021,00**	**35.786,00**	**29.531,00**	**0,00**	**0,00**	**91.338,00**
Equipamento Equipment	2.700,00	6.000,00	5.250,00	0,00	0,00	**13.950,00**
TOTAL	**28.721,00**	**41.786,00**	**34.781,00**	**0,00**	**0,00**	**105.288,00**

Plano de financiamento
Finance plan

DESCRIÇÃO DESCRIPTION	2007	2008	2009	2010	2011	TOTAL
Financiamento solicitado à FCT Requested funding	28.721,00	41.786,00	34.781,00	0,00	0,00	**105.288,00**
Financiamento próprio Own funding	0,00	0,00	0,00	0,00	0,00	**0,00**
Outro financiamento público Other public-sector funding	0,00	0,00	0,00	0,00	0,00	**0,00**
Outro financiamento privado Other private funding	0,00	0,00	0,00	0,00	0,00	**0,00**
Total do Projecto **Total of the project**	**28.721,00**	**41.786,00**	**34.781,00**	**0,00**	**0,00**	**105.288,00**

4. Justificação do orçamento
4. Budget justification

4.1. Justificação dos recursos humanos
4.1. Human resources justification

Tipo	Nº de pessoas	Duração	Custo envolvido (€)
(BI) Bolsa de Investigação (Lic. ou Bacharel)	2	12	17880

Justificação

Two students are expected to take part in the project research while doing their Msc work, as described in the tasks.

The additional costs cover the mandatory Seguro Social Voluntário (~ 78 euro/month) and insurance against personal accidents (~ 300 euro/year), per person.

Tipo	Nº de pessoas	Duração	Custo envolvido (€)
(BI) Bolsa de Investigação (Mestre)	1	36	35280

Justificação

One student will take part in the project as the central topic of his PhD activities. The expected contribution is described in the tasks. Alternatively, more than one student may be involved after finishing their MSc thesis.

The additional costs cover the mandatory Seguro Social Voluntário (~ 78 euro/month) and insurance against personal accidents (~ 300 euro/year).

Custo total: 53160

4.2. Justificação de missões
4.2. Mission justification

Tipo	Local	Nº de deslocações	Custo envolvido (€)
Participação em congressos	TBD - international	5	10000

Justificação

The funding covers an estimated total of 5 international missions of one researcher in the course of the project (1 + 2 + 2) for presenting the results of the project. The total cost presented is based on an average cost of 2000 euro for typical Europe and America missions (5 x 2000).

Tipo	Local	Nº de deslocações	Custo envolvido (€)
Participação em congressos	Portugal	3	1200

Justificação

Funding concerning an estimated total of 3 domestic missions of one researcher (1 per year) for presenting the results of the project (3 x 400 euro).

Custo total: 11200

4.3. Justificação de consultores
4.3. Consultants justification
(vazio)
(void)
4.4. Justificação de aquisição de serviços e manutenção
4.4. Acquisition of services and maintenance justification

Tipo	Custo envolvido (€)
Payment for small tasks	2000

Justificação

This concerns funding requested in order to pay small tasks, like maintaining the public and private web sites for the project, small programming and testing tasks and the like.

Custo total: 2000

4.5. Justificação de outras despesas correntes
4.5. Current expenses justification

Tipo de despesa	Custo envolvido (€)
Small expenses	1250

Justificação

This funding covers small, assorted expenses, such as bibliography (books, reviews), software, blank media, printing posters for project presentation, etc.

Custo total: 1250

4.6. Justificação do Equipamento
4.6. Equipment justification

4.6.1. Equipamento já disponível para a execução do projecto
4.6.1 Available equipment

Tipo de equipamento	Fabricante	Modelo	Ano
Laptops/desktops	Various	Various	2004

4.6.2. Discriminação do equipamento a adquirir
4.6.2. List of new equipment requested

Tipo de equipamento	Fabricante	Modelo	Custo envolvido (€)
Laptops	TBD	TBD	5250

Justificação

The project will use computers assigned to the research and teaching staff of the host instituition. However, some of this equipment should be replaced in the course of the project. Equipment is also needed to support the participant students. For this purpose, we intend to acquire at least 3 laptops (3 x 1750 euro).

PCs	TBD	TBD	2700

Justificação

We intend to acquire up to 3 desktop PCs (3 x 900 euro) to be used by members of the project (mainly students).

Server	TBD	TBD	6000

Justificação

We intend to acquire one server-class equipment, with some number crunching and networking capabilities, to assist in the simulation and prototype validation tasks.

Custo total: 13950

5. Equipa de investigação
5. Research team

5.1 Lista de membros
5.1. Members list

Nome **Name**	**Função** **Role**	**Grau académico** **Academic degree**	**%tempo** **%time**
Sérgio Marco Duarte	Inv. Responsável	DOUTORAMENTO	30
José Augusto Legatheaux Martins	Investigador	AGREGAÇÃO	15
Margarida Paula Neves Mamede	Investigador	DOUTORAMENTO	10
Nuno Manuel Ribeiro Preguiça	Investigador	DOUTORAMENTO	5

Total: 4

5.2. Lista de membros a contratar durante a execução do projecto
5.2. Members list to hire during project's execution

Membro da equipa **Team member**	**Função** **Role**	**Duração** **Duration**	**%tempo** **%time**
(BI) Bolseiro de Investigação (Lic. ou Bacharel) 2	Bolseiro	12	100
(BI) Bolseiro de Investigação (Lic. ou Bacharel) 3	Bolseiro	12	100
(BI) Bolseiro de Investigação (Mestre) 1	Bolseiro	36	100

Total: 3

6. Projectos financiados
6. Funded projects

Lista de projectos financiados
Funded projects list

Referência	**Título**	**Estado**
POSC/EIA/59064/2004	FEW - Files EveryWhere	Em curso

Total: 1
(Os detalhes de cada projectos estão disponíveis clicando na referência correspondente)
(Details for each project are available by clicking on the corresponding reference)

7. Indicadores previstos
7. Expected indicators

Indicadores de realização previstos para o projecto

Expected output indicators

DESCRIÇÃO DESCRIPTION	2007	2008	2009	2010	2011	**Total**
A - Publicações **Publications**						
Livros Books	0	0	0	0	0	**0**
Artigos em revistas internacionais Papers in international journals	0	0	1	0	0	**1**
Artigos em revistas nacionais Papers in national journals	0	0	0	0	0	**0**
B - Comunicações **Communications**						
Comunicações em encontros científicos internacionais Communications in international meetings	1	2	2	0	0	**5**
Comunicações em encontros científicos nacionais Communications in national meetings	0	0	0	0	0	**0**
C - Relatórios **Reports**	1	1	1	0	0	**3**
D - Organização de seminários e conferências **Organization of seminars and conferences**	0	0	0	0	0	**0**
E - Formação avançada **Advanced training**						
Teses de Doutoramento PhD theses	0	0	1	0	0	**1**
Teses de Mestrado Master theses	0	0	2	0	0	**2**
Outras Others	0	0	0	0	0	**0**
F - Modelos **Models**	0	1	0	0	0	**1**
G - Aplicações computacionais **Software**	0	0	1	0	0	**1**
H - Instalações piloto **Pilot plants**	0	0	0	0	0	**0**
I - Protótipos laboratoriais **Prototypes**	0	0	1	0	0	**1**
J - Patentes **Patents**	0	0	0	0	0	**0**
L - Outros **Other**						
	0	0	0	0	0	**0**
	0	0	0	0	0	**0**
	0	0	0	0	0	**0**

Acções de divulgação da actividade científica
Scientific activity spreading actions
(Estimula-se a apresentação e propostas neste âmbito que possibilitem a aproximação da actividade científica ao grande público)
(It is strongly desired the presentation of proposals within this subject that will approach science to the general public.)
This project will develop a P2P application for the Dissemination of Web Syndication content.

The prototype will be freely available, and its potential exposure is the public in general.

The P2P concept has received a tremendous amount of bad publicity for being traditionally associated with copyright violations. Yet, P2P is being heavily researched and is recognized by the scientific community as a powerful tool for creating innovative distributed applications.

LiveFeeds will target free Web Syndication content. As such, it has the potential to be a showcase for legal uses of P2P techniques. This could help raise the awareness of the general population regarding this and other promising technologies facing legal problems.

---

8. Anexo técnico
8. Technical addendum

–

8.1. Resumo
8.1. Abstract

Resumo (em português)
Abstract (in portuguese)

Este projecto tem como pano de fundo a "Web Syndication": um conceito que enriquece a WWW ao permitir aos utilizadores manterem-se ao corrente do conteúdo dos seus "sites"e "blogs"favoritos e do conteúdo da Web em geral. A Web Syndication está a revelar-se apelativa, como atesta o número considerável de sítios de topo que já oferecem fontes RSS [1]. De facto, a capacidade de agregar e actualizar automaticamente um grande número de resumos de conteúdos Web tornam-na muito prática e uma extensão natural à WWW, especialmente considerando que o modelo de base desta última pressupõe que a iniciativa de sondar periodicamente os conteúdos está a cargo dos próprios utilizadores.\nA seu tempo, a evolução da Web Syndication poderá levá-la a tornar-se num autêntico serviço de notificação para a WWW. Apesar de tal ser algo

especulativo, um levantamento informal dos conteúdos RSS disponibilizados por provedores de primeira linha mostra que esta já está a ser usada para disseminar eventos populares com requisitos temporais estreitos, tais como alertas noticiosos e resultados desportivos. Esta utilização é apelativa mas levanta dúvidas quanto a eventuais problemas de escala, os quais se agudizam bastante se for exigida uma maior actualidade dos eventos. A raíz do problema assenta no facto de que a Web Syndication é explorada, correntemente, com base em modelos [1, 2] que no essencial descrevem os conteúdos e que, para efeitos práticos, ainda confiam que os clientes façam consultas periódicas (polling) para detectar actualizações. Este modelo simples teve a vantagem permitir uma implantação rápida e incremental do serviço sobre a infraestrutura Web já existente mas, no futuro, poderá levar a um considerável esforço suplementar nos gastos de largura de banda tanto pelos provedores como pelos clientes. Basicamente, melhorar a frescura da informação através de consultas periódicas é em si mesmo caro e quando realizado em maior escala traduz-se tendencialmente num problema realmente dispendioso e intratável. No limite, tal poderia levar a uma oposição e abandono da Web Syndication, repetindo o colapso do PointCast [3]. \nOutras questões levantam-se com as abordagens actuais à Web Syndication. Uma delas respeita à sua desadequação face à mobilidade. De novo, delegar nos clientes a consulta periódica é na generalidade dos casos uma má escolha para ambientes móveis, levando ao desperdício de energia e largura de banda, os quais são frequentemente escassos e caros. Em cenários orientados à mobilidade, onde a desconexão e baixa conectividade são norma, o utilizador estará pouco interessado em receber um relato integral de cada uma das fontes que subscreve e, em vez disso, preferirá ser alertado para eventos particularmente importantes. De facto, a possibilidade de filtrar os eventos e dar precedência a alguns deles é parece igualmente útil e desejável para a generalidade dos utilizadores. Atender estas questões traz desafios interessantes que se traduzem principalmente num problema de notificação em larga-escala com requisitos de filtragem, o qual constitui um tema de investigação actual [4, 5].\n \nÉ nossa intenção atacar os problemas discutidos acima através de uma solução baseada na cooperação dos clientes. Do ponto de vista do servidor nada mudaria e seriam os clientes com interesses semelhantes que actuariam concertadamente de modo a atingir os objectivos desejados. Juntos, estariam envolvidos em recolher, filtrar e distribuir os conteúdos

RSS entre eles à medida do necessário. Para tal, serão investigadas formas de introduzir princípios P2P como base da infraestrutura de disseminação/notificação de eventos, a qual será modelada de acordo com um modelo realmente "push"[3], em alternativa ao modelo actual. Para além disso, o projecto procurará melhorar o modelo da Web Syndication com funcionalidades relacionadas com a filtragem e precedência de eventos compatíveis com os objectivos de larga escala e, especificamente, sem dar aos servidores essa pesada incumbência. É ainda do nosso interesse explorar conceitos das redes sociais, para que utilizadores com interesses semelhantes (subscrições) possam partilhar fontes de informação e beneficiar da sua experiência como um colectivo. Esta é uma oportunidade que se perde no modelo actual, onde os utilizadores interagem com os servidores isoladamente. Com a mudança que propomos para um modelo de infraestrutura de disseminação cooperativa baseada em princípios P2P ela torna-se viável como também natural. Existe, no entanto, algum perigo de abuso associado a esta mudança. Por isso, este aspecto tem que ser tratado na proposta com as salvaguardas devidas para evitar problemas semelhantes com as mensagens não solicitadas de correio electrónico. Nesse sentido, será importante assegurar que mudança de um modelo isolado de subscrição de conteúdos para um modelo cooperativo e completamente distribuído não irá fazer perigar outras necessidades de segurança básicas, como a privacidade, autenticidade e um uso justo.

Resumo (em inglês)

Abstract (in english)

LiveFeeds focuses on Web Syndication - a concept that enhances the World Wide Web experience by helping users to keep track of the updates made to their favorite websites, blogs and web content in general. Web Syndication is attractive and becoming increasingly popular, as testified by the number of top sites that already offer RSS [1] feeds. In effect, the ability to automatically aggregate and update many web feeds, summarizing the content of many different sources, makes Web Syndication a convenient and natural addition to the World Wide Web, especially since the underlying model of the latter has relied implicitly on the users to poll the web sources themselves in order to stay up-to-date.\nIn time, Web Syndication could very well evolve into a full fledged notification service for the World Wide Web. Although that is still somewhat speculative, an informal survey of the syndication content provided by major web content providers already shows that it is being used to disseminate popular, time-critical events, such as news alerts and sport results. Such usage is surely appealing but raises important scalability concerns, which are worsened if more aggressive requirements in liveliness (event freshness) are factored in. The root of the problem lies in the fact that Web Syndication is currently being pursued based on models [1, 2] that essentially only detail the web feed structure and, for all practical purposes, still rely on client-side polling as the basis for collecting updates. This straightforward model had the advantage of allowing a quick, incremental deployment of the service on top of the already existing server infrastructure but, in the future, it could lead to considerable strains on the communication budgets of content providers and users alike. Improving liveliness is expensive using polling techniques and multiplied by large-scale it translates to the costly problem of dealing with large numbers of users polling frequently and systematically many web feeds. This could ultimately lead to a backlash or complete abandonment of Web Syndication, reenacting a similar episode of PointCast ́s debacle [3].\nThere are other issues in the current approaches to Web Syndication. An important one pertains to its poor suitability when considered in conjunction with Mobility. Again, client-side polling is, generally, a poor choice for mobile environments. It is wasteful in terms of energy and bandwidth, both of which are often scarce and expensive in those situations. In mobility-driven scenarios, where disconnection and poor connectivity are the norm, the user is likely less interested in a full report of each of the web feeds he/she subscribes but, instead, would prefer to be alerted for particularly important events. In fact, it is arguable that the ability to filter web feeds and prioritize events is desirable even for the generality of users. Addressing these issues introduces interesting challenges to Web Syndication, mainly in the form of large-scale notification with filtering requirements, which is a current research topic, e.g. [4, 5].\nWe intend to address the problems discussed above by pursuing large-scale Web Syndication in a cooperative manner. Under such principle, the server-side model will be preserved and it will be up to the clients with similar interests to act in concert towards the desired goal. Together, they will be involved in collecting, filtering and distributing web feeds content among themselves. In order to achieve these goals, we will devise ways to incorporate P2P principles to form the basis of an alternative Web Syndication event dissemination infrastructure, which will be modeled primarily around a (true) push-based model [3]. Additionally, the project will seek ways to enhance the Web Syndication model with support for fine-grained filtering and prioritization of syndication content in ways compatible to its large-scale dissemination goals and, in particular, without cumbering the servers with that additional effort.\nWe also want to explore social networking concepts to allow users with similar interests (subscriptions) to share information automatically and benefit from their experience as a collective. This opportunity is missed in the prevalent Web Syndication model, in which users act alone and interact with servers completely isolated from each other, yet it becomes feasible and natural with the move to a P2P-based infrastructure, as we advocate. However, this can be a potential source for abuse, so this aspect must be addressed in the proposed solution in a suggestive, non-intrusive spirit and include proper safeguards to avoid spam-like problems. In addition to that, it will be important to ensure that the move from an individualistic subscription model to a distributed, cooperative one does not jeopardize other basic security requirements, such as privacy, authenticity and fair use.

8.2. Objectivos

8.2. Objectives

Descrição dos Objectivos do Projecto

Project Objectives (description)

Project Objectives (description)

The goal of the project is to design and prototype a scalable system for Web Syndication content notification. The solution must adhere to the existing Web Syndication standards and use client-side cooperation to minimize server-side involvement. \nTo this end, we need to:\n * Develop novel large-scale dissemination algorithms that incorporate fine-grained filtering to improve overall efficiency and reduce network usage waste. Develop accurate experimental models to drive the simulation, evaluation and refinement of aforementioned algorithms;\n * Provide support for disconnection and poor connectivity of mobile devices. The aim is to allow mobile clients to subscribe web syndication content and observe graceful service degradation;\n * Implement, validate and distribute a prototype aggregator application based on the proposed algorithms.

Descrição dos Objectivos do Investigador Responsável

Principal Investigator Objectives (description)

The researcher is a member of the Large-Scale Distributed Computing Systems research group and his research interests lie on event notification platforms and large-scale content dissemination in general. \n \nThe goal of this project is to design a scalable system for Web Syndication content notification. As such, the objectives of the project match closely the research interests of the PI and, thus, provide a context for the continuation of this line of work.

8.3. Estado da Arte

8.3. State of the Art

Descrição do Estado da Arte

State of the Art (description)

Content dissemination has proved to be a hard problem in the Internet. This is so because the Internet is not truly a broadcast medium and lacks wide support for multicast communication (see [6]). The tremendous success of the Internet rests on top of a point-to-point messaging capability. Performing content dissemination directly on top of that means that a separate copy of the data has to be sent to each receiver. The bandwidth requirements and the load on the servers of this naive multicast technique makes this approach unfeasible for large audiences. To make it work in the Internet of today, transparent client redirection, backed by mass-replication, is used by virtually all the big players in the World Wide Web, who recruit the services of specialized companies, such as Akamai Technologies, at a considerable monetary cost.\nThe development of overlay networks, where message routing is performed at the application-level between processes, was used as a workaround for the lack of multicast. Several examples of early overlay multicasting exist ([7, 8, 9]). These solutions, however, did not truly address the scalability issues. By performing routing at an higher level of abstraction, these overlay networks face the problem of having little or no information regarding the underlying topology of the physical network; so, to optimize the network parameters, they need to invest bandwidth to sample the network and perform housekeeping tasks, which ruins scalability.\n \nThe real breakthrough regarding scalability of overlay networks is tied to the advent of the first generation of DHTs (Distributed Hash Tables) and is credited to [10, 11, 12]. These DHTs were the first that could claim to scale to millions of nodes. Distributed applications, including P2P multicasting (see [13, 14]) could now target the Internet. There was a catch, though. DHTs are structured overlay networks that achieve their efficiency by exploiting mathematical topologies that bear little or no relation to the underlying physical network. This leads to very efficient routing in the logical DHT domain and low overheads, but the observed communication latency is far from optimal because nodes that are neighbors in the logical domain can be scattered across the entire Internet. This and other problems have made DHTs a very hot research topic to this day and a steady stream of improvements have been proposed, including DHTs offering improved performance and even better theoretical limits, such as [15, 16]. For more pointers on structured overlays and their issues, see [17].\nEpidemic or gossip-based dissemination has also been proposed as solution to the scalability problem [18, 19, 20]. The basic idea is simple and consists in having data dissemination mimic the spread of an infectious disease. In practice, the overlay peers form a random graph with a high degree of link redundancy and, often, do not have full knowledge of the network membership. This, indeed, allows the network to scale to a degree, but there are several problems. The same problems that plague DHTs can also affect epidemic networks. The random graph can be a poor approximation of the physical network, leading to considerable delays. Additionally, the guarantees that messages will be delivered to all recipients are probabilistic and there is a marked tradeoff between performance and communication overhead. Still, despite its flaws, epidemic dissemination has found practical use in the Internet, as proved by Bittorrent [21], a file sharing protocol that has become popular and responsible for a sizable fraction of the Internet traffic. Despite being largely epidemic, Bittorrent still depends somewhat on a central server to bootstrap (and loosely maintain) the network.\nAdding filtering to large-scale data dissemination raises the problem to another dimension, which is shared with a related area of research, known as publish/subscribe systems. \nPublish/subscribe systems, also known as event notification platforms, offer a communication model in which data, commonly referred as events, produced by the publishers is to be delivered to each of the subscribers if it matches and is accepted by a previously supplied filter. Initially, the research effort concerning events was focused on the description of events and filters, and the search for fast algorithms for detecting matches. Many aspects relating to distribution were left unsolved; centralized implementations or distributed ones that required manual administration of the support servers were common (e.g. [22, 23]. Naturally, all these these platforms were aimed at small-scale environments.\nContributions from overlay networks research and, soon after DHTs, to publish/subscribe systems added improved distribution models with self-organization and self-repairing capabilities and allowed notification platforms to resort to overlay multicast to target much wider scopes, including the Internet (e.g. [24, 25, 26, 27]). Nevertheless, despite de advances, solutions for exploiting filters to save bandwidth, by discarding early unwanted events, are still far from perfect and the problem is still being researched on.\nFilter-aware data and event dissemination has trivial solutions but they are generally undesirable. If filtering is performed by the subscribers, the overhead of filter evaluation is minimal at the expense of message waste from unwanted events being maximal. On the other hand, if event filtering is performed at the source, message waste is minimized but the source has to do all the work and the solution will not scale very well. Ideally, one would organize the dissemination network to reflect the coverage of the filters according relationships among them, but general solutions to achieve this are not obvious and the type of events and filters used plays an important role. Some of the latest attempts to solve this problem [4] try to use a structured overlay (DHT) to store the filters and then use incoming events as queries to retrieve the relevant filters that need to be evaluated. But, it is has proved to be rather difficult to map multi-dimensional events onto the single key spaces that most DHTs use. One reasonable approach is to deal with each dimension individually, but then the critical problem of performing efficient multiple joins still remains. A related problem, which is performing range queries on multidimensional values stored in DHTs, which could provide valuable insights, is also facing similar difficulties (see [28] for more details).\nReferences\n[1] RSS 2.0 Specification. http://blogs.law.harvard.edu/tech/rss.\n[2] M. Nottingham and R. Sayre, The Atom Syndication Format. RFC 4287, 2005. \n[3] M. Franklin and S. Zdonik, Data in your face : push technology in perspective, in SIGMOD ’98, 1998.\n[4] I. Aekaterinidis and P. Triantafillou, Pastrystrings: A comprehensive content-based publish/subscribe dht network, in ICDCS ’06, 2006.\n[5] E. Anceaume, et al., A semantic overlay for self- peer-to-peer publish/subscribe, in ICDCS ’06, 2006.\n[6] H. W. Holbrook and D. R. Cheriton, IP Multicast Channels: EXPRESS Support for Large-scale Single-source Applications, in SIGCOMM ’99, 1999.\n[7] J. Jannotti, et al., Overcast: Reliable Multicasting with an Overlay Network, in \nOSDI ’00, 2000.\n[8] Y.-H. Chu, S. G. Rao and H. Zhang, A Case for End System Multicast, in SIGMETRICS ’00, 2000.\n[9] D. Pendarakis, et al., ALMI: An Application Level Multicast Infrastructure, in USITS ’01, 2001.\n[10] S. Ratnasamy, et al., A Scalable Content-addressable Network, in SIGCOMM ’01, 2001.\n \n[11] I. Stoica, et al., Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications, in SIGCOMM ’01, 2001. \n[12] A. Rowstron and P. Druschel, Pastry: Scalable, Decentralized Object address, and Routing for Large-Scale Peer-to-Peer Systems, in Middleware 2001, 2001.\n[13] S. Ratnasamy, et al., Application-Level Multicast

address, and Routing for Large-Scale Peer-to-Peer Systems, in Middleware 2001, 2001.\n[13] S. Rathasamy, et al., Application-Level Multicast Using Content-Addressable Networks, in Networked Group Communication, 2001.\n[14] S. Q. Zhuang, B. Y. Zhao, and et al., Bayeux: An

Architecture for Scalable and Fault-tolerant Wide-area Data Dissemination, in ACM Workshop on Network and Operating Systems Support for Digital Audio and Video, 2001.\n[15] D. Malkhi, M. Naor, and D. Ratajczak, Viceroy: a Scalable and Dynamic Emulation of the Butterfly, in PODC '02, 2002.\n[16] M. F. Kaashoek and D. R. Karger, Koorde: A Simple Degree-optimal Distributed Hash table, in IPTPS '03, 2003.\n[17] S. El-Ansary and S. Haridi, An overview of structured overlay networks, in Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless and Peer-to-Peer Networks, 2005. \n[18] P. Eugster, et al., Lightweight Probabilistic Broadcast, in International Conference on Dependable Systems and Networks (DNS 2001), 2001.\n[19] I. Gupta, et al., Efficient Epidemic-style Protocols for Reliable and Scalable Multicast, in IEEE International Symposium on Reliable Distributed Systems, 2002. \n[20] A.-M. Kermarrec, et al., Probabilistic Reliable Dissemination in Large-Scale Systems, IEEE TPDS '03, 2003.\n[21] Bram Cohen, Incentives build robustness in bittorrent, 2003.\n[22] A. Carzaniga, D. S. Rosenblum, et al., Design and Evaluation of a Wide-area Event Notification Service," ACM Transactions on Computer Systems, 2001.\n[23] D. S. Rosenblum and A. L. Wolf, A Design Framework for Internet-Scale Event Observation and Notification, in European Software Engineering Conference (ESEC/FSE 97), 1997. \n[24] Sérgio Duarte, DEEDS - A Distributed and Extensible Event Dissemination Service. PhD thesis, Universidade Nova de Lisboa, 2005. \n[25] A. I. T. Rowstron and et al., SCRIBE: The design of a large-scale event notification infrastructure, in Networked Group Communication 2001, 2001. \n[26] L. F. Cabrera, M. B. Jones, and et al., Herald: Achieving a Global Event Notification Service, in 8th Workshop on Hot Topics in Operating Systems, 2001. \n[27] P. Pietzuch and J. Bacon, Hermes: A Distributed Event-Based Middleware Architecture, in DEBS '02, 2002. \n[28] A. Datta, et al., Range queries in trie-structured overlays, in IEEE P2P '05, 2005.

8.4. Resultados e Repercussões
8.4. Results and Repercussions

Divulgação de Resultados (descrição)
Diffusion of Results (description)

Dissemination of results will be achieved mainly in two ways: \nscientific publication in peer reviewed media and software distribuition.\nResults will be expected following the first year of the project, and will include: \n \n- publication of papers in established international journals, \nconferences and workshops with scientific refereeing (such as \n IEEE Int. Conf. on Dist. Computing Systems (ICDCS); ACM Symp. on Operating Systems Principles (SOSP); Usenix Symp. on Operating Systems Design and Implementation (OSDI); IEEE Int. Symp. on Reliable Dist. Systems (SRDS); International Conference on Distributed Event-Based Systems (DEBS)); \n \n- publication of papers in national conferences (CRC); \n \n- availability of technical reports and prototypes in the site \ncreated for the project; \n \n- contribution to new dissertations (MSc and PhD).

Repercussões (descrição)
Repercussions (description)

This project will provide a context for the advanced schooling of new \npostgraduate students/researchers (MSc and PhD). \n \nAdditionally, the project will contribute with techniques that we expect\nwill be useful in addressing related problems.\nFinally, we expect the project to produce practical results, in the form of an application\nsoftware that will be made available to a wide audience.

8.5. Regionalização
8.5. Regionalization

Região Region	Percentagem Percent
Norte	0
Centro	0
Lisboa e Vale do Tejo	100
Alentejo	0
Algarve	0
Região Autónoma dos Açores	0
Região Autónoma da Madeira	0

Descrição
Description

The project targets the Internet as whole, accordingly the knowlege and software it will produce are general in nature and do not hold any regional bias.

8.6. Tarefas
8.6. Tasks

Lista de tarefas (5)
Task list (5)

Designação da tarefa Task denomination	Data de início Start date	Data de fim End date	Pessoas * mês Person * months
T2 - Models & Simulation Environment	01-01-2007	31-12-2007	14
T3 - Dissemination Algorithms	01-01-2007	31-12-2008	21
T1 - Management & Dissemination of Results	01-01-2007	31-12-2009	22
T4 - Prototype Development	01-06-2008	31-05-2009	17
T5 - Prototype Evaluation & Refinement	01-01-2009	31-12-2009	8

(Os detalhes de cada tarefa estão disponíveis clicando na designação correspondente)
(Details for each task are available by clicking on the corresponding denomination)

8.7. Referências Bibliográficas
8.7. Bibliographic references

Ano Year	Publicação Publication

2005	Sérgio Duarte, "DEEDS - A Distributed and Extensible Event Dissemination Service ". PhD thesis, Universidade Nova de Lisboa, 2005.
2003	Sérgio Duarte, José Legatheaux Martins, Henrique J. Domingos, and Nuno Preguiça, "A case study on event dissemination in an active overlay network environment., " in DEBS (H.-A. Jacobsen, ed.), ACM, 2003.
2005	Nuno Preguiça, Carlos Baquero, José Legatheaux Martins, Marc Shapiro, Paulo Sérg io Almeida, Henrique João Domingos, Victor Fonte and Sérgio Duarte, FEW: File Ma nagement for Portable Devices," in Proc. of "The International Workshop on Soft ware for Portable Storage (IWSSPS),", Califórnia, USA. March 2005.
2006	Nuno Preguiça, José Legatheaux Martins, Henrique J. Domingos, Sérgio Duarte. Sup porting Multi-Synchronous Groupware: Data Management Problems and a Solution, In ternational Journal of Cooperative Information Systems, Volume 15, Number 2, Jun e 2006.
2005	Margarida Mamede, "Recursive lists of clusters: A dynamic data structure for ran ge queries in metric spaces.," in ISCIS, vol. 3733 of Lecture Notes in Computer Science, pp. 843–853, Springer, 2005.

8.8. Artigos Anteriores
8.8._ Previous Articles

Ano **Year**	**Artigo (endereço na *Internet* - URL)** **Paper (Link in the Internet - URL)**
2005	http://asc.di.fct.unl.pt/~smd/livefeeds/ref1.pdf
2003	http://asc.di.fct.unl.pt/~smd/livefeeds/ref2.pdf
2005	http://asc.di.fct.unl.pt/~smd/livefeeds/ref3.pdf
2006	http://asc.di.fct.unl.pt/~smd/livefeeds/ref4.pdf
2005	http://asc.di.fct.unl.pt/~smd/livefeeds/ref5.pdf

---

9. Ficheiros Anexos
9. Attachments

–

(vazio)
(void)