

ALTO Working Group
Internet-Draft
Intended status: Informational
Expires: January 12, 2012

Z. Dulinski
Jagiellonian University
P. Wydrych
R. Stankiewicz
AGH University of Science and
Technology
July 11, 2011

Inter-ALTO Communication Problem Statement
draft-dulinski-alto-inter-problem-statement-01

Abstract

This draft considers an approach to the optimization of the traffic generated by the overlay (peer-to-peer) applications, where some information on inter-AS (Autonomous System) paths is obtained with the usage of dedicated communication scheme known as inter-ALTO communication.

The large amount of network traffic generated by overlay applications requires effective management. This traffic traverses inter-AS links and thus generates substantial costs for the operators and poses problems with overload on the external and internal links. The traffic is not time-stable as the peers connect and disconnect very often. Additionally, when the overlay traffic is observed on inter-AS links, it can be seen that sources and destinations change in a very short period of time. The ALTO (Application-Layer Traffic Optimization) service provides the information that enables more efficient management of the overlay traffic. Such applications can use the information to perform better-than-random peer selection. The ALTO servers are responsible for a pre-selection procedure; the final selection is done by overlay clients and then the ALTO protocol conveys network information to applications. To be credible, this information should be as close to real network situation as possible. However, some types of data are not held by an operator, but they should be gained on the basis of the additional information exchange with other AS operators. This document presents rationale for the need of introduction of the inter-ALTO communication.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-

Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 12, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Definitions	5
3. The Problem and Motivation	6
3.1. Route Asymmetry	8
3.2. Many ASes within One ISP	8
3.3. Different Types of Business Relations	9
3.4. Congestion Avoidance	9
3.5. Proximity Awareness	9
3.6. Remote ISP's Preference	10
3.7. Coordination of ISPs' Policies	10
3.8. Sensitivity of Topology Information	11
3.9. Outdated Information	11
3.10. Mobile Networks	11
3.11. Route Aggregation	12
4. Usage of the Mechanisms Offered by the ALTO Protocol	12
5. Security Considerations	13
6. IANA Considerations	14
7. Acknowledgements	14
8. References	15
8.1. Normative References	15
8.2. Informative References	15
Authors' Addresses	15

1. Introduction

This document describes the rationale for a communication to be used between ALTO servers located in different autonomous systems (AS). Such an inter-ALTO communication extends the ALTO service [RFC5693] capabilities and provides additional information on remote peers, i.e., peers located in other ASes. To make the consideration more clear we distinguish local AS and remote ASes. Local AS is the one from which perspective we describe the communication. Local peers are located in the local AS and are served by a local ALTO server. On the contrary, all other peers are located in remote ASes. Those peers are referred to as remote and are served by remote ALTO server. This basic terminology adheres to majority of considerations in this document.

The motivation for the ALTO service as discussed in the ALTO problem statement [RFC5693] focuses on the overlay traffic optimization based on information gathered from the Internet Service Provider (ISP) domain, i.e., an Autonomous System (AS). Due to the suggested approach, information on the AS internal topology and some routing information obtained from the global Internet (the BGP routing tables) may be used for the peer selection procedures. The data transfer cost can be also incorporated. However, there are some parameters which can be used for the better peer selection mechanism, but they are not available in the local AS and must be obtained from outside, i.e., from a remote AS. For example, the BGP routing information available in the AS identifies only the upstream traffic. The information about the downstream traffic is not present or is incomplete and the full BGP information for this traffic could be obtained from the remote AS containing the subnetwork where the peer sending downstream traffic is attached. In order to obtain such data, we propose the inter-ALTO communication.

It is assumed that the ALTO servers are deployed in the local and remote ASes. The ALTO server located in the client AS can request desired information from the ALTO server which is located in the remote AS. Each server is managed by a respective ISP. Each ISP decides what type of information can be exposed to the requesting party. The ALTO server responds with the type of information that was previously agreed to exchange. Each local ALTO server has to discover the address of the remote ALTO server before starting the communication. The discovery procedure may use the IP addresses of remote peers for the establishment of IP addresses of remote ALTO servers. The detailed analysis of this functionality is out of scope of this document.

The information delivered by remote ALTO servers may be used by a local ALTO server to perform advanced rating/ranking procedure of

peers. The general idea is as follows.

1. A peer receives a list of other peers from a tracker, i.e., a list of potential candidates to communicate with.
2. A peer uses the ALTO protocol [I-D.ietf-alto-protocol] to send the list of peers to its local ALTO server.
3. Local ALTO server obtains additional information on remote peers by communicating respective remote ALTO servers. If sufficient information is available locally and the inter-ALTO communication is not needed, this step is omitted.
4. Using ISP specific policies and values of parameters associated with remote peers the local ALTO server performs rating/ranking procedure.
5. Sorted/rated list of peers is sent back to the peer with usage of the ALTO protocol.

The requirements, syntax and detailed operation of the inter-ALTO communication as well as the rating/ranking procedure is out of scope of this document.

2. Definitions

In the scope of this document we use all the definitions specified in the Section 2 of ALTO problem statement [RFC5693]. Moreover, the following terms have the special meaning.

Local Peer: A peer which belongs to the same Autonomous System to which the ALTO client belongs.

Remote Peer: A peer which belongs to another Autonomous System than the one to which the ALTO client belongs.

Local AS: The Autonomous System to which the ALTO client belongs.

Remote AS: An Autonomous System to which a remote peer belongs.

Local ALTO Server: The ALTO server serving the ALTO client and the local peers.

Remote ALTO Server: An ALTO server serving remote peers.

3. The Problem and Motivation

ALTO server optimization capabilities are limited by the fact that they use information available locally only. It can be shown that more information on remote peers, a routing path, or remote ISP preferences would be useful. The data from remote ASes obtained by the external interface as shown in Figure 1 of the ALTO protocol draft [I-D.ietf-alto-protocol] may have a substantial significance for the management of overlay traffic (e.g. with respect to peer rating, ranking, or the choice of the best peers). The suggested approach to deliver these types of information is defined in the inter-ALTO communication discussed in this document.

The ALTO service may also be used by the client-server applications, supporting the best choice of the mirror servers. If some mirror servers are in other ASes than the client's AS, some information about the network conditions in the remote ASes may be delivered only by the ALTO servers localized in these ASes. Both clients and mirror servers may communicate with their local ALTO servers for delivering traffic optimization parameters. As an ALTO client communicates only with its local ALTO server, the information from remote ASes is accessible only via inter-ALTO communication.

The ALTO-based traffic optimization may be also used in the context of the Content Delivery Networks (CDNs) [I-D.jenkins-alto-cdn-use-cases]. The draft by Niven-Jenkins et al. shows how CDNs may benefit from the information provided by the ALTO protocol. Local information, however, may be not sufficient to optimize the request routing process. The information gathered from ALTO servers in other domains may be needed.

The basic ALTO architecture presented in Figure 1 of the ALTO protocol draft [I-D.ietf-alto-protocol] defines the external interface used to communicate with other information sources like remote ALTO servers. The ALTO Protocol draft, however, does not define what information should be exchanged between ALTO servers to optimize the traffic. The inter-ALTO communication proposed by the current document implements the external interface as defined by the ALTO protocol draft.

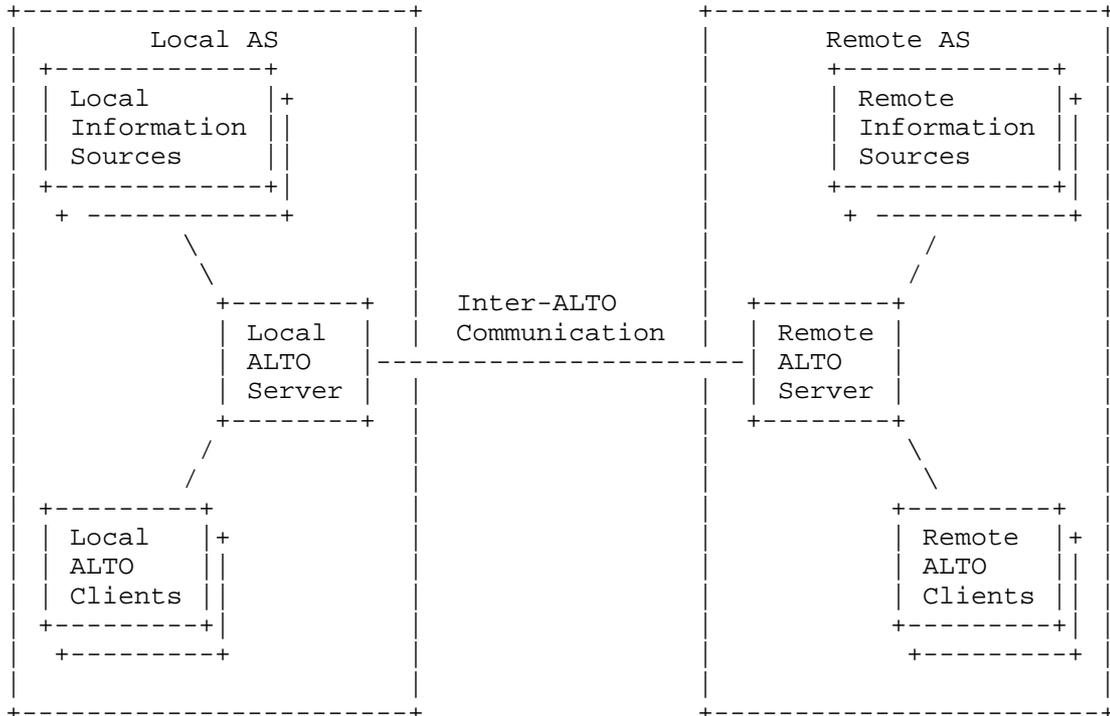


Figure 1: Inter-ALTO communication architecture.

The architecture of the Inter-ALTO communication is shown in Figure 1. Both ALTO servers gather the information from their information sources like routing protocols, provisioning policies, or dynamic network information sources. The local ALTO server needs to communicate with a remote ALTO server to obtain information which is available only at the entities in the remote AS.

In particular, the following key aspects motivate the proposal of the inter-ALTO communication.

- o Route asymmetry.
- o Different types of business relations.
- o Congestion avoidance.
- o Proximity awareness (distance to the remote AS), e.g.:
 - * number of inter-AS hops;

- * delay (RTT).
- o Remote ISP's preference.
- o Coordination of ISPs' policies.
- o Outdated information.

3.1. Route Asymmetry

The communication between two ASes does not need to follow the same path in the upstream and downstream direction. It was shown that about 29% of paths between AS pairs in the Internet are fully symmetric, i.e., upstream and downstream traffic follows exactly the same path [ICC.optimal]. In 51% of cases the number of inter-AS hops is different for the upstream and downstream direction. Additionally, in 50.5% of all path pairs a neighbor AS for upstream and downstream paths are different.

The ALTO server can obtain routing information locally (e.g. from BGP routers) and can determine the upstream path. Information about the downstream path is usually not easily available. Some additional routing information can be obtained from Looking Glass Servers, but not all ASes provide them. The inter-ALTO communication provides the ability to exchange the relevant information between ALTO servers. Especially, the downstream path can be reliably determined using the information provided by remote ALTO server. In the light of route asymmetry in the Internet such information appears to be necessary for a better optimization of a peer rating/ranking algorithm, as assumption that the inter-AS routes follow symmetrical paths can give not only sub-optimal, but misleading and, in effect, harmful results.

3.2. Many ASes within One ISP

An ISP may possess a complex topology network composed of many autonomous systems. Current ALTO specification allows for deployment of independent ALTO servers in each AS. In such a case the overlay traffic management performed by the ALTO server is restricted to a single AS since cost maps have a local meaning. An ISP operating a multi-AS network may be interested in managing the traffic in the whole administrative domain in a consistent and coordinated manner. The information possessed by a single ALTO server is insufficient. To obtain a complete knowledge on the multi-AS network a communication between ALTO servers is needed. As a result, local cost maps originating from different autonomous systems may be coordinated. A uniform cost map reflecting the whole network structure may be created and distributed between ALTO servers.

3.3. Different Types of Business Relations

Two basic business relations between ISPs may be distinguished.

When two ISPs agree to exchange the traffic without any charge, such a relation is called peering. The inter-domain link between the respective ASes is also called a peering link. Usually, there is no charge if the difference between traffic volumes passing such a link in different directions does not exceed a previously agreed limit.

The other case occurs when one ISP serves as a network provider to another ISP (e.g. relation between tier 2 and tier 3 ISPs). In such a case one ISP (acting as a customer) has to pay the other ISP (acting as provider) for the traffic sent over the inter-AS link connecting them. The real monetary cost of the traffic volume exchanged on such a link depends on agreements between ISPs. In general, some links may be considered as cheaper or more expensive.

AS may be connected to many other ASes with various agreements. The cost of the inter-AS traffic transfer may differ depending on which neighbor AS the path passes. For this reason an ISP may prefer that its own customers exchange data with remote peers located in such ASes that the path directed to them passes cheaper links. The ALTO server may sort peers taking into account these criteria. To receive almost complete information on routing paths to and from different remote domains the information provided by remote ALTO server using inter-AS communication may be helpful.

3.4. Congestion Avoidance

A peer rating/ranking procedure may also take into account the congestions on inter-AS links. An ISP is able to monitor queues on its inter-domain links and assign metrics indicating the buffer occupancy or bandwidth utilization. These metrics can express percentage use of buffers or bandwidth on a particular inter-AS link. If one inter-domain link is congested it is desirable to promote peers reachable through lightly loaded links. Again, information provided by the remote ALTO server would support such optimization. The aim of the inter-ALTO communication is not to replace the existing congestion avoidance mechanisms. The idea is to support the present mechanism by the exchange of parameters describing the load on the inter-AS links.

3.5. Proximity Awareness

For a set of reasons (e.g. the performance of an application) the ALTO server may suggest its customers to connect to remote peers located in its proximity. The simplest measure of proximity is the

number of inter-AS hops. As indicated above, due to the route asymmetry, the number of hops may significantly differ between the upstream and downstream paths. Such information for the downstream path may be provided by the remote ALTO server. A more advanced metric of proximity can be found in the delay that can be approximated by exchanging messages between ALTO servers. The ALTO servers can be equipped with an application-layer ping functionality which only operates between ALTO servers. By exchanging special packets prepared by the ALTO servers, these servers can estimate delay and packet loss.

3.6. Remote ISP's Preference

If two ISPs agree on a cooperation, the remote ALTO server may provide its preference parameters (remote preference parameters) indicating which peers are better from the point of view of the remote ISP. For instance, the AS in which the remote ALTO server is located may possess two subnetworks connected to the operator's core network by distinct links. It may happen that a connection to one of the subnetworks is cheaper than the other. The remote operator may prefer connections through cheaper link, so peers located in the subnetwork transferring data via this cheaper link are preferred.

The remote preference parameter may be also used when a remote ISP wants to suggest peers which are connected to the Internet through access links of higher capacity. This way, the remote ALTO server, without exposing the exact values of access link bandwidth, may indicate peers with higher throughput. The remote preference parameters have only local meaning, i.e., their values are comparable for peers located in the same AS only.

If a remote ISP does not want to reveal numerical values of network parameters related to its peers (such information might be considered as confidential) the remote ALTO server may perform a rating/ranking procedure and assign priority parameter to its peers. The rating/ranking criteria may remain hidden for the requesting local ALTO server.

3.7. Coordination of ISPs' Policies

Operators may have an incentive to coordinate their efforts in order to decrease transfer costs on inter-AS links or improve quality experienced by peers, i.e., coordinate their peer rating/ranking strategies. This way, operators may avoid contradictory strategies resulting in inefficiency of rating/ranking algorithms. Operators may agree to promote each other's peers.

For example, it may happen that operator A wanting to decrease

traffic on one of its links discourages its own peers from communicating with peers located in operator B's domain. On the other hand, operator B would consider peers located in a domain of operator A as very attractive for its own peers. As a result, rating/ranking procedures performed by respective ALTO servers give contradictory results what may decrease the effectiveness of these procedures. To avoid such a situation, the inter-ALTO communication is needed.

Another example of a usefulness of coordination of policies is clustering of ASes. Recent studies [IJNM.unfairness] have shown that locality promotion might be ineffective or even harmful if used in AS with small number of peers. A proposed solution is to create a cluster of two or more ASes. Then ALTO servers serving different ASes in the cluster treat all peers located in the cluster as if they were in a single AS. In other words, from a point of view of locality promotion algorithm all peers located in the cluster are local, regardless of their home AS.

3.8. Sensitivity of Topology Information

The minimum information that the remote AS provides to the local ALTO server via the inter-ALTO communication may be the number of inter-AS hops and the number of the local AS's neighbor in the downstream path (the full downstream AS_PATH may be not exchanged). Such information does not reveal any sensitive information neither on the ISP internal topology details nor remote AS connections with other ASes, but does provide basic and very useful information for the local ALTO server.

3.9. Outdated Information

It is expected that some information (parameters) from routing protocols that will be used in the rating/ranking procedures may outdate. Also information related to the network performance is constantly changing. Therefore, the information obtained from the remote AS requires updates. This updates may be generated on request (pull mechanism), on event base schema or periodically (push mechanism). The inter-ALTO communication should be equipped with mechanisms for updates. The need for the present information describing network conditions and some routing parameters are arguments for introducing specific protocol for the communication between ALTO servers.

3.10. Mobile Networks

The inter-ALTO communication may be very useful for mobile network operators and content providers serving mobile clients. An ALTO server may recognize mobile clients and properly assign them to PIDs.

Some information about the mobile network resources gathered from mobile network nodes located in different networks should be exchanged between operators for better than random peer selection. ALTO servers should possess information which allows to make proper peer selection, taking into account, e.g., the mobile network load (including the load in the radio access network and in the circuit- and packet-switched domains).

After collecting the load information, the ALTO server may assign priorities. These priorities may exemplify the load in some parts of the radio access network. Via the inter-ALTO communication, the priorities may be passed to the other operator's networks where other clients are located. Relying on this information, the ALTO server may optimize the connections between clients.

3.11. Route Aggregation

The BGP protocol allows the aggregation of specific routes into one route. In such a case the aggregate route is advertised. The full path is either lost completely or the AS set information is available. In the latter case only the set of ASes behind the aggregating router is known but the detailed information about the routing path, including AS sequence and AS-hop count, is lost. From the overlay traffic optimization point of view the knowledge on ASes located behind aggregating router and the number as well as sequence of inter-AS hops may be useful, e.g., because of route asymmetry problem described earlier (Section 3.1). The solution for this problem is information exchange between ALTO servers located in ASes ahead and behind the router aggregating routes.

4. Usage of the Mechanisms Offered by the ALTO Protocol

The basic ALTO protocol architecture allows an ALTO server to communicate with a third party through the external interface. The inter-ALTO communication may use some functionalities offered by the ALTO protocol [I-D.ietf-alto-protocol].

Server Information Service: This service defined by the ALTO protocol may be extended in order to provide information about server's ability to cooperate with other ALTO servers. Thanks to this service, the other ALTO servers may acquire the information about available parameters and their definitions. These parameters may be used by cooperating ALTO servers for the peer rating/ranking procedures. The access for this service may be restricted. Some information may be accessible only by the privileged ALTO servers after the successful authentication.

ALTO Information Services: These services has been defined to provide the query information services for ALTO clients. All the information delivered by these services has local meaning. This information is related to the locally defined parameters describing a particular ISP's network. Some part of this information managed by a remote ALTO server may be useful for the requesting local ALTO server. The requesting ALTO server obtains this information via inter-ALTO communication. After receiving the response, the local ALTO server has to perform some calculations, scaling, merging, or adaptation of the received parameters. In this way the local ALTO server may conform to both its internal network topology and measurements, and the external ones. However, it should be stressed that the ALTO Information Services is designed for communication between ALTO clients and servers, not for the inter-ALTO communication.

Network Map: This structure is defined by the ISP and reflects the internal structure of the ISP network. This structure has only a local meaning and, generally, it is not unique for all entities within the Internet. A particular network map can be used by different operators. The requesting ALTO server usually has to perform some prediction of the external topology on its own. The ALTO server has to apply its own rules and definitions. The PIDs, defined in the remote ALTO server, have to be mapped on the PID structure defined in the local AS.

Cost Map: This structure also has the local meaning. The local ALTO server may receive the network map and the cost map from a remote ALTO server. These costs may require recalculation in order to unify the cost measures in the local AS. After these operations, if it is needed, the rating/ranking procedure can be performed.

5. Security Considerations

The communication between ALTO servers requires authentication and authorization procedures. In some cases it may require establishment of the secured tunnels between the partner ALTO servers. The minimum security requirements for the inter-ALTO communication is out of scope of this document.

The inter-ALTO communication allows ALTO servers to exchange any parameters which improve the performance of the overlay traffic, or, generally, allows them to manage overlay traffic. In order to achieve this results a group ISP may exchange sensitive data, the exchanged parameters may be confidential. They should not be

accessible by a third party, e.g., some other ISPs or peers.

An ISP may have its own policy how organize the overlay traffic and this policy may use a number of parameters during the evaluation procedure. The policy result may be delivered to peers in many ways. It can take the form of a sorted peer list without any parameters, a sorted list with some parameters which are derived from the parameters exchanged in the inter-ALTO communication, or raw exchanged parameters. ISPs may have an incentive not to expose these parameters in the raw form to peers. The mentioned sensitive parameters require applying a higher level of the security procedures.

In order to keep the exchanged parameters confidential it may be reasonable to keep the communications between peers and ALTO server from communication among ALTO servers by the protocol differentiation separated. Different security procedures may be easier to manage if these communication procedures take the form of two distinct protocols. This protocol separation allows to define mechanisms which are specific for the inter-ALTO communication only. The protocol should not allow to use this mechanism by overlay peers. The set of procedures for the inter-ALTO communication is expected to be separated from the client ALTO communication and this can be achieved by distinct protocols.

6. IANA Considerations

This document has no actions for IANA.

7. Acknowledgements

The authors would like to thank following people for the valuable discussions (in alphabetical order):

- o Piotr Cholda (AGH University of Science and Technology)
- o Mirosław Kantor (AGH University of Science and Technology)
- o Jan Medved (Juniper)
- o Stefano Previdi (Cisco)
- o Robert Varga (Pantheon)

This work has been partially supported by the EU through the ICT FP7 Project SmoothIT (FP7-2007-ICT-216259).

8. References

8.1. Normative References

- [I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol",
draft-ietf-alto-protocol-09 (work in progress), June 2011.
- [ICC.optimal]
Dulinski, Z., Kantor, M., Krzysztofek, W., Stankiewicz,
R., and P. Cholda, "Optimal Choice of Peers based on BGP
Information", Proceedings of 2010 IEEE International
Conference on Communications (ICC), May 2010.
- [RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic
Optimization (ALTO) Problem Statement", RFC 5693,
October 2009.

8.2. Informative References

- [I-D.jenkins-alto-cdn-use-cases]
Niven-Jenkins, B., Watson, G., Bitar, N., Medved, J., and
S. Previdi, "Use Cases for ALTO within CDNs",
draft-jenkins-alto-cdn-use-cases-01 (work in progress),
June 2011.
- [IJNM.unfairness]
Lehrieder, F., Oechsner, S., Hossfeld, T., Staehle, D.,
Despotovic, Z., Kellerer, W., and M. Michel, "Mitigating
unfairness in locality-aware peer-to-peer networks",
International Journal of Network Management, Volume 21,
Issue 1, pp. 3-20, January/February 2011.

Authors' Addresses

Zbigniew Dulinski
Jagiellonian University
ul. Reymonta 4
Krakow 30-059
Poland

Phone: +48 12 663 5571
Fax: +48 12 633 4079
Email: dulinski@th.if.uj.edu.pl

Piotr Wydrych
AGH University of Science and Technology
al. Mickiewicza 30
Krakow 30-059
Poland

Phone: +48 12 617 4817
Fax: +48 12 634 2372
Email: piotr.wydrych@agh.edu.pl

Rafal Stankiewicz
AGH University of Science and Technology
al. Mickiewicza 30
Krakow 30-059
Poland

Phone: +48 12 617 4036
Fax: +48 12 634 2372
Email: rstankie@agh.edu.pl

