



IPv6 Connectivity Deployment in Sofia University Network

Vesselin Kolev, Georgi Naidenov, Stefan Dimitrov
Sofia University

Difficulties in Allocating IPv6 Address Space for an End User

- Currently the Regional Internet Registrar (RIR) RIPE accepts only LIRs' (Local Internet Registrar) requests for allocation of IPv6 address space:
<http://www.ripe.net/ripe/docs/ripe-421.html>
- Sofia University (SU) to acquire a LIR status would not be justified neither financially nor functionally, having in mind its status in the Internet community.
- 32-bit network mask address space is extremely large and as a LIR SU should be obliged to allocate at least 200 48-bit network mask IPv6 address sublocations to its customers.
- RIR RIPE does not allocate IPv6 spaces to end customers (as it does with IPv4) due to the lack of accepted IPv6 "PI (Provider Independent)" allocations regional policy. At present just a proposal is issued in this respect:
<http://ripe.net/ripe/policies/proposals/2006-02.html>



IPv6 Address Space Request through LIR

- It is more difficult for end users in Bulgaria to be allocated IPv6 address space because most of the Bulgarian Internet providers (ISPs), that have LIR status, are not still well aware of IPv6 and have no such connectivity.
- At the time when SU technical team began its efforts to find way to allocate IPv6 address space the Academic Network (NREN) was the only LIR with IPv6 connectivity.
- NREN to allocate to SU a part of its /32 segment would have been the easiest way.
- But such endeavour would not benefit Bulgarian Internet space: to urge commercial ISPs to invest in new technologies as IPv6.
- The technology would remain closed among academic bodies, not to benefit society as a whole.
- So, collaborative approach was taken: commercial ISP with LIR status to have desire to deploy IPv6.



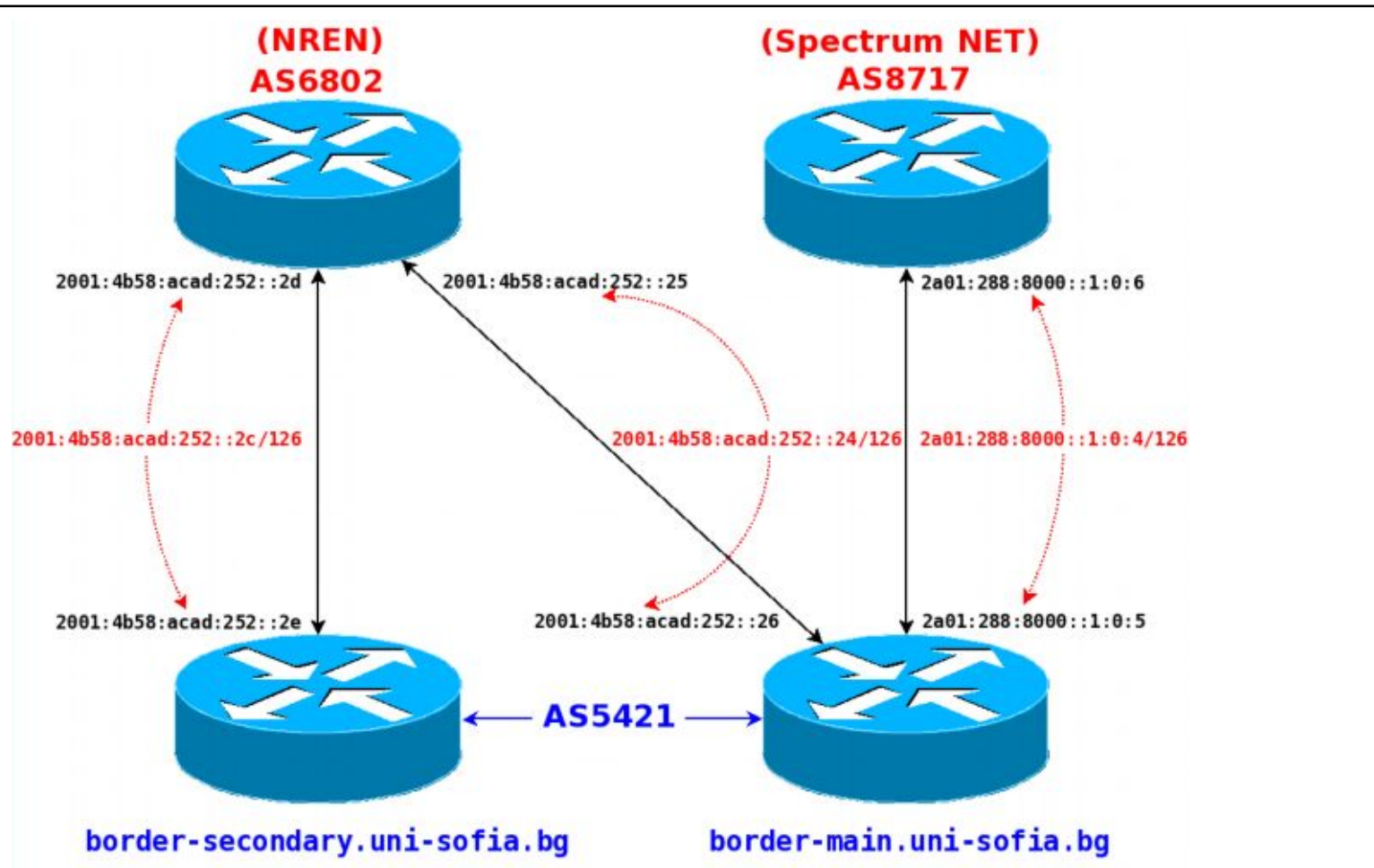
IPv6 Address Space Request through LIR (cont.)

- LIR "Spectrum NET" was convinced to take from RIR RIPE its /32 IPv6 address segment and to set its network equipment for transit of IPv6 traffic.
- "Spectrum NET" allocated address subspace to SU and in this way became the first Bulgarian ISP with IPv6 connectivity and transit.
- That is a precedent how the academic community appears as a catalyst for the adoption of a new technology in mass usage.
- The following address block was allocated by "Spectrum NET":
 - 2a01:288::/32
- The block status is ALLOCATED-BY-RIR. SU was delivered the block:
 - 2a01:288:8000::/35 with status ALLOCATED-BY-LIR.

Organizing the IPv6 traffic transit to/from SU autonomous system (AS5421)

- IPv6 traffic transit to/from AS5421 border routers is accomplished by three BGP4+ sessions.
- Two of them are with AS6802 (NREN) border router, and the third with AS8717 (Spectrum NET) border router.
- border-secondary.uni-sofia.bg border router has only one BGP4+ session with AS6802, and border-main.uni-sofia.bg has two BGP4+ sessions: one with AS6802 and the second with AS8717 border router.
- Next are illustrated the three BGP4+ sessions for IPv6 traffic transit:

BGP4+ transit sessions between AS5421 routers and routers of AS6802 and AS8717





Organizing the IPv6 traffic transit to/from AS5421 (cont.)

- An agreement was reached with NREN for IPv6 traffic transit through the transmission media of the European Research Network GEANT2.
- A similar agreement was reached for IPv6 traffic transit through the AS of Spectrum NET on MAN transmission media.



Organizing the IPv6 traffic transit to/from AS5421 (cont.)

- AS5421 border routers are BGP originator of the IPv6 prefix 2a01:288:8000::/35 in the general BGP Internet table.
- In order SU prefix to be generated through AS6802 border routers, and from there to GEANT2 ones and to the rest of the Internet, redefinition of AS6802 and AS20965 border routers' filters was needed.
- This task was accomplished by NREN technical team.



Organizing the IPv6 traffic transit to/from AS5421 (cont.)

- It was agreed with Spectrum NET to generate SU prefix 2a01:288:8000::/35 with no aggregation.
- In principle Spectrum NET should aggregate 2a01:288:8000::/35 prefix to 2a01:288::/32, but in this particular case 2a01:288:8000::/35 is transferred through two different providers and no aggregation is needed.
- Otherwise the input traffic to SU IPv6 space would pass solely through NREN, 2a01:288:8000::/35 prefix is more specific than 2a01:288::/32.



Organizing the IPv6 traffic transit to/from AS5421 (cont.)

In conclusion, the process of generation looks like:

- Generating through Spectrum NET:

2a01:288::/32

2a01:288:8000::/35

- Generating through NREN and GEANT2:

2a01:288:8000::/35



Some Difficulties in Generating 2a01:288:8000::/35 Prefix

- Some border routers do not accept 2a01:288:8000::/35 prefix, but accept the aggregated one: 2a01:288::/32.
- For instance, the border routers of AS 1280, belonging to Internet Systems Consortium (ISC).
- It was agreed with their technical staff to correct the prefix filters.
- But a general question arises: are operators of border routers in Internet inclined at all to pass prefixes with network mask different from /32 and /48.



Some difficulties in Generating 2a01:288:8000::/35 Prefix (cont.)

- According RIR RIPE documents, accepted by the other RIRs, the smallest IPv6 prefix not to be filtered should be with 48-bit network mask, i.e the minimal allocation delivered by LIR to end user.
- So, filtering of prefixes with mask between 32 and 48 bits is not a good practise and is not agreed around the Community.
- Such filters could lead to concentration of transit traffic at some points, even to limited visibility of some IPv6 address spaces, hence deterioration in transit service.
- As some RIRs have already accepted PI allocations (e.g. AfriNIC) and accomplish direct /48 allocations and greater (e.g. /47, /46), RIR RIPE is expected to do the same soon.
- So no filters should be present in BGP4+ transit sessions' descriptions, rejecting prefixes with network mask between 32 and 48 bits.



IPv6 Addresses Delivery Policy to Independent Units in SU

- Each independent unit within SU can issue a request to University Computer Centre (UCC) for IPv6 address space.
- The request is served in case it is possible to transfer IPv6 traffic to that unit and address space with 48-bit network mask is allocated.
- A SU unit can be allocated an IPv6 segment by default in case new routers are installed at its site, directly connected to SU backbone network.
- By 2009 all independent units in SU should run IPv6.



IPv6 Addresses Delivery Policy to Independent Units in SU (cont.)

At present all the independent units at campus "Lozenets" possess IPv6 address spaces, delivered from the 2a01:288:8000::/35 prefix:

- UCC:
2a01:288:8003::/48
- Faculty of Mathematics and Informatics (FMI):
2a01:288:8002::/48
- Faculty of Physics (Phys):
2a01:288:8005::/48
- Faculty of Chemistry (Chem):
2a01:288:8004::/48
- Laboratory of Chemical Physics and Engineering (LCPE):
2a01:288:8001::/48



IPv6 Addresses Delivery Policy to Independent Units in SU (cont.)

- In all the above units address allocation in the internal networks is done by Radvd according RFC2461.
- That is, the user has to do nothing about address and gateway settings. It is done automatically.
- In Windows XP IPv6 addressing has to be activated.
- In Windows Vista the setting is activated by default.
- All Linux or BSD distributions should have IPv6 addressing activated by default.

Address Spce Reserved for Management and Control

- Part of IPv6 address space is assigned to network control devices and the basic support servers.
- The 2a01:288:8000::/48 address space is assigned as follows:
 - 2a01:288:8000::/96 to the routers in SU backbone;
 - 2a01:288:8000:aaaa::/96 to the basic servers implementing common network services for SU.



BGP4+ Reflector Scheme and OSPF6 area 62.44.127.0

- The border routers of AS5421 establish BGP4+ sessions to the routers of the independent units.
- For redundancy two BGP4+ sessions are maintained.
- SU network implements good practises concerning dynamic routing, so neither static nor default routes are used.
- In all, the routing process in SU, even within the internal network, is based on BGP.



BGP4+ Reflector Scheme and OSPF6 area 62.44.127.0 (cont.)

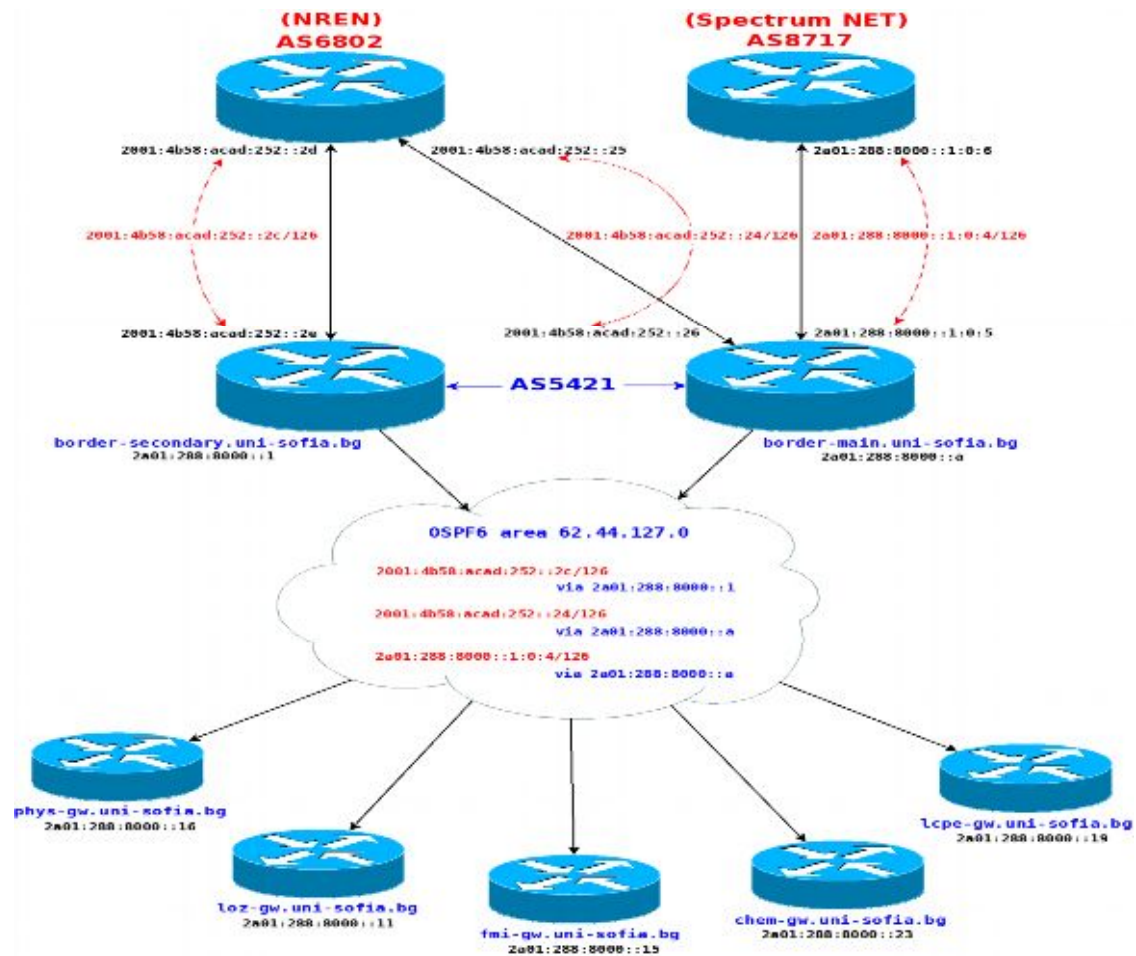
- Each of the independent unit's router receives by the BGP4+ sessions the full BGP tables of AS5421 border routers.
- But, the independent units' routers to be able to forward output traffic on iBGP, they need additional information to reach the "next hop" through the transit autonomous systems.
- The OSPFv3 protocol is used for the above purpose, and OSPF6 area 62.44.127.0 in particular.



BGP4+ Reflector Scheme and OSPF6 area 62.44.127.0 (cont.)

- By OSPF6 area 62.44.127.0 the recursive problem concerning output traffic is solved on the basis of the routes delivered to the BGP4+ session.
- OSPF6 area 62.44.127.0 points the direction for each "next hop", the IP of the border router of the AS that has advertised the route. This "next hop" is not in one and the same Ethernet segment with backbone router.
- All stated above is illustrated below (the routes to be advertised are shown in the cloud):

The Structure of OSPF6 area 62.44.127.0





BGP4+ Reflector Scheme

- BGP4+ reflector is introduced in order to have optimal and flexible internal routing policy implemented.
- The aim of the reflector scheme is to set the shortest path when routing within an AS of routers.
- Each router turns out to be an originator of the network it serves as a gateway. And it advertises the prefix of this network to both border routers of AS5421, the latter acting as "reflectors".
- The "next hop" parameter in the BGP prefix is in the same Ethernet segment, and there is direct exchange of traffic between the networks behind the backbone routers bypassing the "reflectors", i.e. the border routers.
- In this way traffic through and local traffic processing load of border routers are reduced significantly.



Delegating ip6.arpa Domains for SU Address Space 2a01:288:8000::/35

- IPv6 address space of SU is not requested directly from RIR RIPE.
- It is used through LIR "Spectrum NET" and can't be registered directly in the zone.
- So, hierarchical approach through the name servers of LIR "Spectrum NET" is taken for the delegation of the ip6.arpa zones of domains for IPv6 address space of SU.



Delegating ip6.arpa Domains for SU Address Space 2a01:288:8000::/35 (cont.)

- On 3 Oct. 2006 RIPE NCC as RIR was authorized to allocate for its members (LIRs) address allocations from 2a00:0000::/12 segment:

<http://www.iana.org/assignments/ipv6-unicast-address-assignments>

- Next are shown the borders of this segment, the prefix length and type:



Information for RIPE 2a00:0000::/12 Segment:

[IPV6 INFO]

Expanded Address	- 2a01:0000:0000:0000:0000:0000:0000:0000
Compressed address	- 2a01::
Subnet prefix (masked)	- 2a00:0:0:0:0:0:0:0/12
Address ID (masked)	- 1:0:0:0:0:0:0:0/12
Prefix address	- fff0:0:0:0:0:0:0:0
Prefix length	- 12
Address type	- Aggregatable Global Unicast Addresses
Network range	- 2a00:0000:0000:0000:0000:0000:0000:0000 - 2a0f:ffff:ffff:ffff:ffff:ffff:ffff:ffff



Delegating ip6.arpa Domains for SU Address Space 2a01:288:8000::/35 (cont.)

- As it was shown above, on top of the hierarchy of ip6.arpa should be the zone of the domain 0.a.2.ip6.arpa, as "2a0" is the constant "prefix" in all address subsegments of 2a00:0000::/12 segment.
- LIR "Spectrum NET" was authorized to control and allocate address space 2a01:288::/32 and in the zone of domain 0.a.2.ip6.arpa delegation of ip6.arpa domain 8.8.2.0.1.0.a.2.ip6.arpa was done, served by name servers ns.spnet.net and purgatory.spnet.net.
- Next follows information for the address segment of LIR "Spectrum NET":



Information for LIR "Spectrum NET" 2a01:288::/32 Segment:

[IPV6 INFO]

Expanded Address	- 2a01:0288:0000:0000:0000:0000:0000:0000
Compressed address	- 2a01:288::
Subnet prefix (masked)	- 2a01:288:0:0:0:0:0:0/32
Address ID (masked)	- 0:0:0:0:0:0:0:0/32
Prefix address	- ffff:ffff:0:0:0:0:0:0
Prefix length	- 32
Address type	- Aggregatable Global Unicast Addresses
Network range	- 2a01:0288:0000:0000:0000:0000:0000:0000 - 2a01:0288:ffff:ffff:ffff:ffff:ffff:ffff



Information for SU 2a01:288:8000::/35 Segment:

[IPV6 INFO]

Expanded Address	- 2a01:0288:8000:0000:0000:0000:0000
Compressed address	- 2a01:288:8000::
Subnet prefix (masked)	- 2a01:288:8000:0:0:0:0/35
Address ID (masked)	- 0:0:0:0:0:0:0/35
Prefix address	- ffff:ffff:e000:0:0:0:0:0
Prefix length	- 35
Address type	- Aggregatable Global Unicast Addresses
Network range	- 2a01:0288:8000:0000:0000:0000:0000:0000 - 2a01:0288:9fff:ffff:ffff:ffff:ffff:ffff



Delegating ip6.arpa Domains for SU Address Space 2a01:288:8000::/35 (cont.)

- In the zone of the domain 8.8.2.0.1.0.a.2.ip6.arpa the domains 8.8.8.2.0.1.0.a.2.ip6.arpa and 9.8.8.2.0.1.0.a.2.ip6.arpa should be delegated.
- The delegation of these domains is according the "glue AAAA" scheme:
- domain 8.8.8.2.0.1.0.a.2.ip6.arpa:
"glue AAAA":
@ORIGIN 8.8.8.2.0.1.0.a.2.ip6.arpa
ns1 A 62.44.96.1
ns2 A 62.44.96.7
ns3 AAAA 2a01:288:8000:aaaa::1
ns4 AAAA 2a01:288:8000:aaaa::7



Delegating ip6.arpa Domains for SU Address Space 2a01:288:8000::/35 (cont.)

- o domain

9.8.8.2.0.1.0.a.2.ip6.arpa:

"glue AAAA":

@ORIGIN

9.8.8.2.0.1.0.a.2.ip6.arpa

ns1 A 62.44.96.1

ns2 A 62.44.96.7

ns3 AAAA 2a01:288:8000:aaaa::1

ns4 AAAA 2a01:288:8000:aaaa::7



DNSSEC Problems with 8.8.8.2.0.1.0.a.2.ip6.arpa and 9.8.8.2.0.1.0.a.2.ip6.arpa

- The contents of the zone of 0.a.2.ip6.arpa domain is electronically signed by RIPE NCC.
- RIPE DB offers automated functions to its members (LIRs) by which they can carry out DNSSEC delegations of their ip6.arpa.
- Then, the LIR as DNSSEC mediator delegates ip6.arpa domains to its customers.



DNSSEC Problems with 8.8.8.2.0.1.0.a.2.ip6.arpa and 9.8.8.2.0.1.0.a.2.ip6.arpa (cont.)


- At present one main obstacle appears, lack of DNSSEC support in PowerDNS software that implements the DNS service at LIR "Spectrum NET".
- We hope that the new version of the software will include DNSSEC support, and the DNSSEC delegation from RIPE shall reach SU.



Collaborating with Register.BG to Introduce "glue AAAA" Delegation for Domains in TLD BG










- At present the domain management form of Register.BG should allow advertising of "glue AAAA" resource records in delegation of domains in TLD BG.
- We express our great thanks to Daniel Kalchev from Register.BG for the support of our idea for IPv6 delegation.

Result of the Collaboration with Register.BG

 **Към Конфигуриране**

Въвеждане/Промяна на сървъри на имена за домейн:

uni-sofia.bg

	име на сървър: ady.uni-sofia.bg	IP адрес: 62.44.96.7	
	име на сървър: ns.digsys.bg	IP адрес: 192.92.129.1	
	име на сървър: ns.uni-sofia.bg	IP адрес: 62.44.96.1	
	име на сървър: ns6.1.uni-sofia.bg	IP адрес: 2a01:288:8000:aaaa::1	
	име на сървър: ns6.2.uni-sofia.bg	IP адрес: 2a01:288:8000:aaaa::7	

Забележка:

1. Задължително е въвеждането на най-малко 2 сървъра на имена.
2. Ако сървъра на имена е в областта на домейна, въвеждането на IP адрес е задължително.
3. Ако сървъра на имена е в област различна от тази на домейна, въвеждането на IP адрес не е необходимо.



Deployment with Open Code Saves a Lot of Money

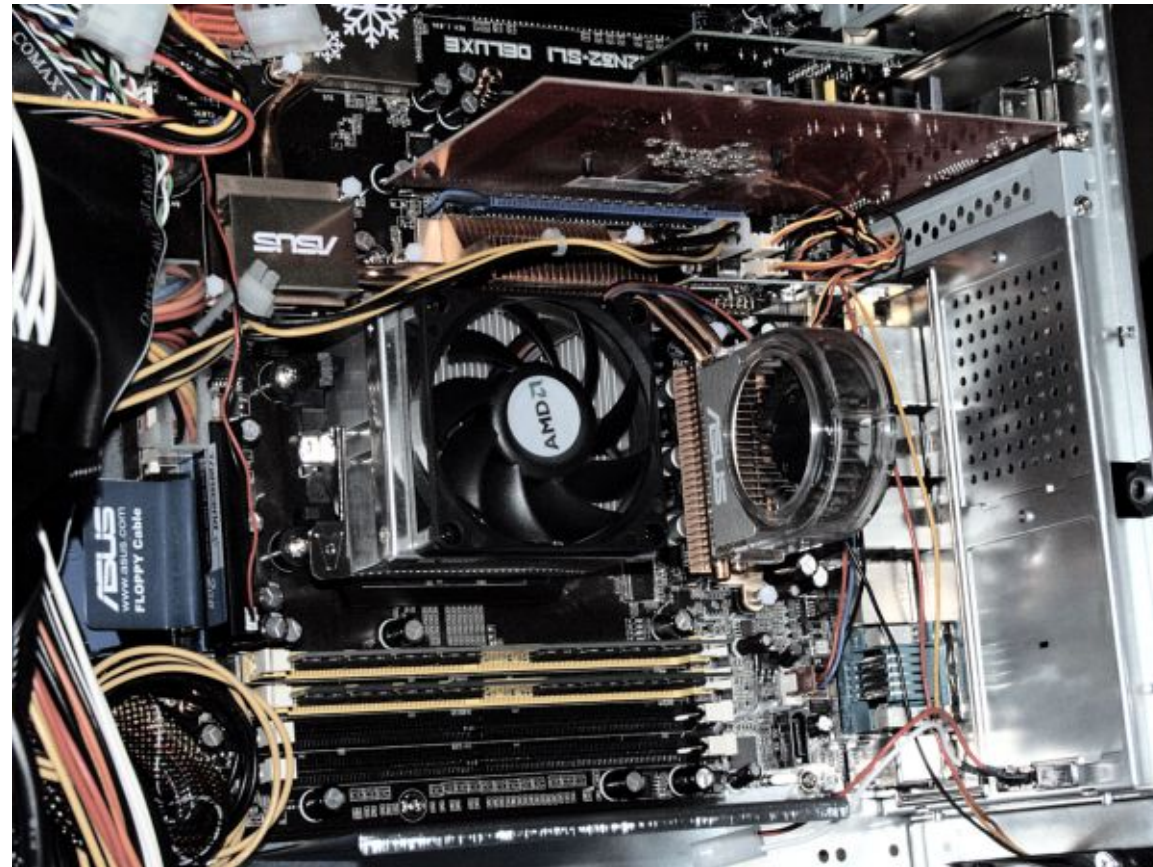
- Border and backbone routers, independent units' routers and the basic servers of SU are implemented on the basis of open code software:
 - Operating system: Linux 2.6;
 - Dynamic routing control software: Quagga.
- By that decision SU saved a great bulk of money achieving the performance and functionality of specialized dual-stack hardware costing tens of thousands of dollars.
- In this way division "Networks and Communications" team at UCC achieved its goal: to introduce the new technology at extremely low price, not compromising performance and functionality.

border-main.uni-sofia.bg

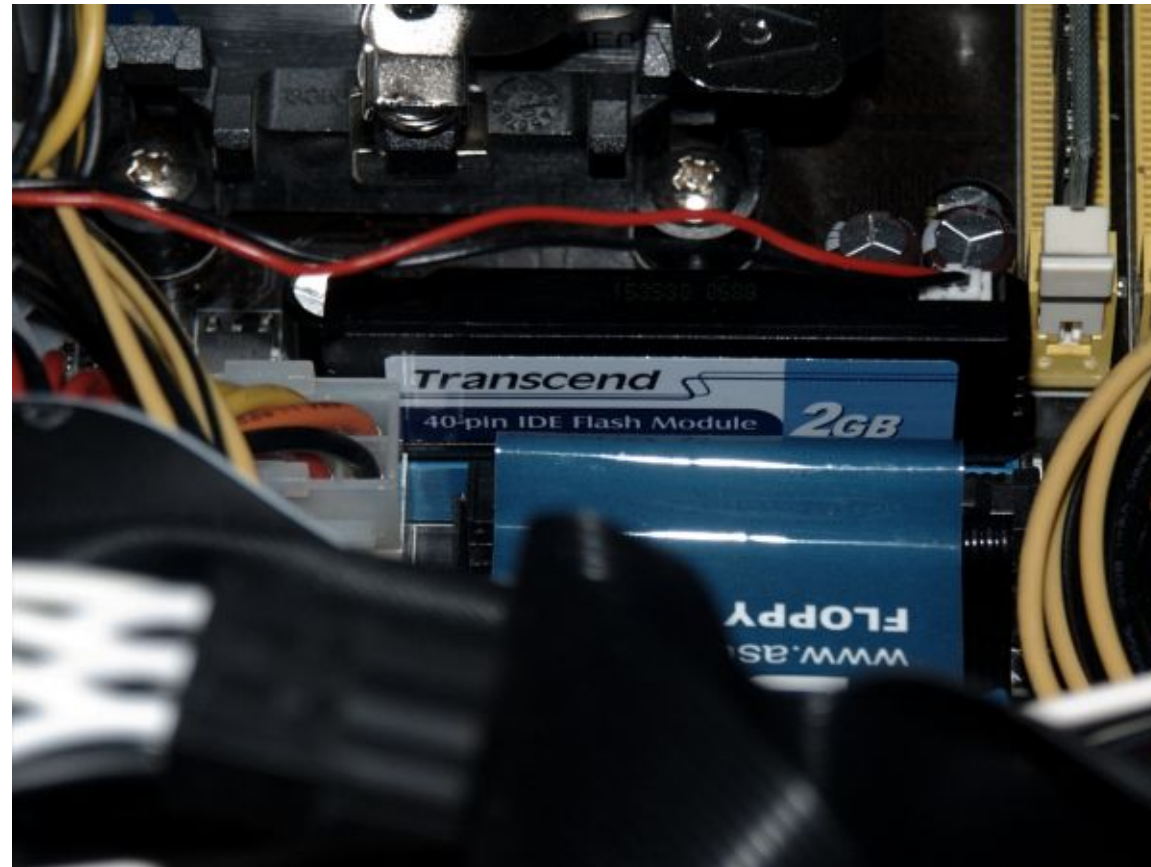
The Box



border-main.uni-sofia.bg
CPU and 2 GB RAM



border-main.uni-sofia.bg
2 GB IDE Flash Module



border-main.uni-sofia.bg

An Inside View

