Seamless Roaming

Adrian Popescu, David Erman, Dragos Ilie, Markus Fiedler, Alex Popescu, Karel de Vogeleer

Blekinge Institute of Technology
Karlskrona, Sweden

October 2008
Outline

- Introduction
- Definition
- Goals and Requirements
- Short History
- Main Challenges
- Types of Handovers
- Standard Bodies
- L2/L3 Handover
- Handover Operations
- IEEE 802.21 MIH
- Internet Mobility
- Mobility Management
- Connectivity Management
- IMS Interworking
- ROVER
- Research Challenges
- Several Important Results
- Conclusions
- References
Introduction

Currently, three important developments in telecom

- Irreversible move towards IP- and SIP-based networking
- Deployment of broadband (wireless) access, e.g., ADSL2+, FTTH, WLAN
- Expansion of mobile communication systems, e.g., UMTS, WLAN, WiMAX

Consequence

- Appearance of more advanced and more bandwidth-demanding applications

New paradigms for the next generation mobile communication

- Always best connected and secured
- E2E seamless service delivery

Handover (HO) has been implemented, so far, within

- Cellular networks
- MIP networks, and
- In media access dependent ways in IEEE 802 networks

Standard bodies: IEEE, 3GPP, 3GPP2, WiMAX, IETF
Introduction (cont’d)

- Traditional HO management was done by using radio specific mechanisms placed at Layer 2 - L2

- Recent research and development based on pushing the HO functionality up to
  - IP layer – L3 (e.g., MIP and FMIP) to easy up the convergence of different technologies
  - Application layer mobility – L5, by using the application protocol SIP

- An important consequence is the need for cross-layer interaction, e.g., between IEEE 802 MAC/PHY and a “roaming” L3

- New solution advanced by BTH: pushing more HO functionality higher up to the application layer – L5
Definition Seamless Roaming

Definition

○ Ability that a user roams in a secure way across different networks while keeping connected and not disturbing ongoing sessions and conversations.

○ Every specific session has own requirements regarding “non disturbance” state with reference to, e.g., error rate, delay, jitter, security, etc.
Goals and Requirements

- Fundamental goals
  - Secured and seamless HO
  - Make the heterogeneous network transparent to the user
  - Design the system architecture such as it is independent of the (wireless) access technology
  - Flexibility

- Other requirements
  - Mobility management: access network location, seamless HO, paging and registration, security provision, policy-based HO
  - Provision of QoS, user and network security, billing, etc.
  - Efficient configuration selection
Short History

- Initial model
  - Develop common standards across IEEE 802 media
  - Define L2 triggers to make FMIP work well
  - Define media independent information to enable cellular/laptop to effectively detect and select networks
  - Define a way to transport this information and these triggers over all IEEE 802 media

- But people wanted cellular inter-working as well; also, wired + wireless was desired with security protection

- Consequence: 802.11 and 802.16 → 802.21

- IMS upcoming

- Need for better prediction mechanisms
Main Challenges

- TCP/IP stack was not designed for mobility but for fixed computer networks
  - Responsibility of individual layers is ill-defined with reference to mobility
  - Consequence: problems in lower layers may create bigger problems in higher layers
  - Higher layer mobility schemes are likely to better suit Internet mobility

- Heterogeneity existent today with reference to
  - Access networks
  - Wireless communication systems
  - Standard bodies
  - Standards
  - Architectural solutions

- Other important problems
  - Lack of interoperability between different types of vendor equipment
  - Lack of standard for handover interfaces
  - Lack of techniques to measure and assess the performance (including security)
  - Incorrect network selection
  - Increasing number of interfaces on devices
  - Presence of different fast handover mechanisms in IETF, e.g., MIPv4, FMIPv6
  - IETF anticipated L2 solutions in standardized form (in the form of triggers, events, etc), but today the situation is that we have NO standards and NOR media independent form
  - Use of L2 predictive trigger mechanisms, which are dependent of L1 and L2 parameters
Types of Handovers

- Horizontal/homogeneous handovers
  - Within single network
  - Localized mobility
  - Limited facilities

- Vertical/heterogeneous handovers
  - Across different networks
  - Global mobility
  - More opportunistic
Handover standards
Handover operation

- HO initiation
- Network and resource discovery
- Network selection
- Network attachment
- Configuration (identifier configuration; registration; authentication and authorization; security association; encryption)
- Media redirection (binding update; media rerouting)
L2/L3 Handover (cont’d)

- Single interface radio
  - Horizontal handover
  - Risk for service disruptions when
    - Performing channel scanning and obtaining QoS information from neighbor PoAs
    - Doing L2 switching and new connection setup, including network entry and route update

- Multiple interface radio
  - Vertical handover
  - No link disconnection during the handover procedure
  - Exchange of L2 frames, with the consequence of risk for large delays
  - Exchange of L3 MIPv6 messages to update route information
L2/L3 Handover (cont’d)

- HO type (horizontal or vertical) and time needed to perform it are determined with the help of
  - Neighbor network information provided by the Base Station (BS)
  - Access Point (AP), and
  - 802.21 Media Independent Handover Function (MIHF) Information Server (IS)

- The Link Going Down (LGD) trigger should be invoked PRIOR to an actual Link Down (LD) event by at least the time required to prepare and to execute a HO procedure

- LGD trigger and prediction
  - Too late LGD trigger – current link may break before a new link is setup
  - Too early LGD trigger – loss of a “working” connection; unnecessary roll-backs of HO cancellations

- Big challenge
  - How to timely generate a LGD trigger that takes into consideration neighboring network conditions and dynamic channel characteristics

- Least Squared Mean (LSM) linear prediction is used to predict expected Link Down (LD) time
L2/L3 Handover (cont’d)

- Main problems L2/L3 handovers
  - Lack of cross-layer interaction between L2 and L3
  - L2 and L3 operate independently of each other
  - Dependence on the limitations of L1, L2, and L3

- FMIPv6 attempts to reduce this problem by using reliable prediction of HO to enable proactive configuration of the involved nodes

- Different MIPv6 versions: Fast MIPv6 (FMIPv6); Hierarchical MIPv6 (HMIPv6); Fast Hierarchical MIPv6 (FHMIPv6)

- Further performance improvements can be obtained by allowing L3 to have control over certain L2 HO related actions

- Conclusion: strong need for further research on cross-layer management!
### Handover Operations

<table>
<thead>
<tr>
<th>HO operation</th>
<th>L2</th>
<th>L3</th>
<th>L5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>Scanning</td>
<td>Router advertisement</td>
<td>Domain advertisement</td>
</tr>
<tr>
<td>Authentication</td>
<td>EAPoL</td>
<td>IKE, PANA</td>
<td>S/MIME</td>
</tr>
<tr>
<td>Security association</td>
<td>802.11i</td>
<td>IPSEC</td>
<td>TLS SRTP</td>
</tr>
<tr>
<td>Configuration</td>
<td>ESSID</td>
<td>DHCP stateless</td>
<td>URI</td>
</tr>
<tr>
<td>Address uniqueness</td>
<td>MAC address</td>
<td>ARP DAD</td>
<td>SIP registration</td>
</tr>
<tr>
<td>Binding update</td>
<td>Cache update</td>
<td>Update CN, HA</td>
<td>SIP re-invite</td>
</tr>
<tr>
<td>Media routing</td>
<td>IAPP</td>
<td>Encapsulation tunneling</td>
<td>Direct media routing</td>
</tr>
</tbody>
</table>
IEEE 802.21 MIH

- **Purpose**
  - Optimize L3 and above handovers
  - Acts across 802 networks and extends to cellular networks (802.3; 802.11; 802.16; cellular)
  - 802.21 MIHF IS server has information about, e.g., location of PoA, list of available networks, cost, L2 information (neighbor maps), higher layer services (ISP, MMS ..)

- **Key benefits**
  - Optimum network selection
  - Seamless roaming
  - Low power operation for multi-radio devices

- **Types of HO**
  - Terminal Controlled
  - Network Initiated, Network Assisted
  - Network Initiated, Network Controlled
IEEE 802.21 MIH (cont’d)

Scope of IEEE 802.21

IEEE 802.21 and IETF
Internet Mobility

- Basic functional requirements for mobility support
  - HO and location management
  - Multi-homing support
  - Support for current services and applications
  - Security

- Limitations of TCP/IP
  - Limitations of Physical and Link Layer (radio channels show limitations compared to fixed networks)
  - Limitations of IP address, it plays the role of both locator and identifier
  - Lack of cross-layer awareness and cooperation
  - Limitations of applications (improper design for mobile environments, e.g., DNS, SIP)
  - Limitations when using different mobility protocols in MN and in network

- Performance metrics relevant for Internet mobility
  - HO latency
  - Packet loss
  - Throughput
  - Signaling
Extending TCP/IP for Mobility

- Mobility support at L3
  - MIPv4; MIPv6; FMIPv6; HMIPv6; FHMIPv6; LIN6; …

- Mobility support at L4
  - Improving TCP performance for mobility: Indirect TCP (I-TCP); Mobile TCP (MTCP)
  - Mobility extension to TCP: TCP Redirection (TCP-R); TCP Migrate (TCP-M); MSOCKS; Mobile UDP (M-UDP); Mobile SCTP (MSCTP); …

- New layer between L3 and L4
  - Host Identity Protocol (HIP); Multiple Address Service for Transport (MAST);

- Mobility support at L5
  - Session Initiation Protocol (SIP); Dynamic Updates in the DNS (DDNS); BTH; …
MIPv4

- Main drawbacks
  - Triangular routing, with risk for large delays
  - Risk for service interruptions due to large delays in HA registration
  - Increased signaling overload

- Suggested improvements
  - Routing optimizations
  - Use of prediction
  - Hierarchical schemes
  - Better paging systems

1. Normal communication between MN and CN
2. MN moves from home to a foreign network
3. MN registers its CoA
4. MN sends packets to CN directly
5. CN sends packets to MN’s home address
6. HA tunnels the packets to MN’s FA
7. FA delivers the received packets to MN
MIPv6

- Two types of L3 mobility
  - Mobile controlled (MC)
  - Network controlled (NC)

- MC demands for mobility stack/client in MN (CMIPv6)
  - MC drawbacks: demand for more resources in MN

- NC demands for networking units in network
  - NC drawbacks: limited mobility domain; use of proxies in the network (PMIPv6)
Basic idea: separation of ID and locator in the IPv6 address
- LIN6 ID is used as node ID
- More tolerant to errors than MIPv4/MIPv6
- Less overhead

LIN6 protocol stack

LIN6 operation

1. Register MN network prefix<->LIN6 ID at bootstrap
2. Query MN FQDN
3. Response MA network prefix+LIN6 ID
4. Query MN LIN6 ID
5. Response MN network prefix
6. Establish connection and start data transfer
7. Move
8. Update the mapping of LIN6 ID<->MN network prefix
9. Go on communication
MSCTP

- Mobile Stream Control Transmission Protocol
- Recently developed IETF transport protocol (RFC 2960)
- Used together with IPSec or Transport Layer Security (TLS) to protect against insecure environments

1. Initiate an association through subnet1
2. Move
3. Add IP into the association
4. Move
5. Change the primary path
6. Move
7. Delete IP from the association
HIP

- Host Identity Protocol
- Designed by the IETF
- Basic idea: separation of location from identity
- Protection against DoS and other security attacks

HIP protocol stack

HIP operation

<table>
<thead>
<tr>
<th>Application layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport layer: &lt;HI, Port&gt; pairs</td>
</tr>
<tr>
<td>Host identity layer: translation between HI and IP address</td>
</tr>
<tr>
<td>Network layer</td>
</tr>
<tr>
<td>Data link layer</td>
</tr>
</tbody>
</table>

1. Register MN HIC<->IP
2. MN FQDN
3. MN HI,RVS IP
4. HI
5. Remainder data
6. Move
7. Update the binding of HI <->IP
8. Go on communication
SIP

- Session Initiation Protocol
- Developed by IETF as an application-layer multimedia signaling protocol (RFC 3261)
- Drawback: risk for HO delay and overload
- Solution: use of prediction
DDNS

- Traditional DNS is restricted in mobile Internet
- DDNS: Dynamic Update of DNS
- Developed by IETF (RFC 2136)
## Functions of Mobility Paradigms

<table>
<thead>
<tr>
<th>Category</th>
<th>Network layer</th>
<th>Transport layer</th>
<th>A new layer</th>
<th>Application layer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIP</td>
<td>L3G6</td>
<td>TCP</td>
<td>UDP</td>
</tr>
<tr>
<td>Handover</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Location</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Multihoming</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Applications</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Security</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
### Required Changes

<table>
<thead>
<tr>
<th>Category</th>
<th>Network layer</th>
<th>Transport layer</th>
<th>A new layer</th>
<th>Application layer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIP</td>
<td>LING</td>
<td>TCP</td>
<td>UDP</td>
</tr>
<tr>
<td>Host</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Routers</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third device</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>TCP/IP layering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[ eurongi logo ]
Mobility Management

- Two major elements
  - Location management
    - HO management
  - Location management
    - Refers to the process used by a network to find out the current attachment point of a mobile user
    - Two phases involved, namely location registration/update and paging

- HO management
  - Refers to the way the network acts to keep mobile users connected when they move and change their position and access points in the network

- Situation today: static algorithms used for Location area (LA) update, no adaptation used to follow the mobility characteristics of the mobile node

- Better performance expected by using dynamic location update mechanisms and paging algorithms

- Basic idea: consider user mobility and accordingly optimize the signaling cost associated with location update and paging

- The goal is to reduce the costs associated with these mechanisms to a minimum
Mobility Management (cont’d)

- Location modeling
  - One- or two- or three-dimensions
  - Levels: location area (controlled by a Mobile Switching Center MSC); cell ID; position inside the cell (geo-location problem)

- Identity of mobile users and the associated billing information is stored in
  - Home Location Register (HLR), respectively
  - Visitor Location Register (VLR)

- Dynamic algorithms for location update
  - Distance-based
  - Time-based
  - Movement-based
  - Movement threshold approach
  - Information theoretic approach

- Mobility modeling and prediction
  - Different criteria: dimension; scale; randomness; geographical constraints; change of parameters; etc
  - Popular models: fluid-flow; random-walk; Gaussian-Markov; geographic-based; group-mobility; kinematic mobility, etc
Connectivity Management

- Increased complexity, mobility refers today more to the change of a logical location with respect to network access point rather than user geographic position.

- Consequence: mobility management becomes more of a connectivity management procedure.

- Two aspects must be considered at vertical HO:
  - HO at device level
  - HO at flow level

- Two general classes of HO mechanisms:
  - Traditional algorithms, with focus on L2/L1 HO
  - Context based algorithms

- Three classes of context based algorithms:
  - Traffic flow based algorithms
  - Simple Additive Weighting (SAW) algorithms
  - Advanced Multiple Criteria Decision Making (MCMD) algorithms

- Another dimension for evaluation and decision:
  - Local
  - Distributed
Example of SAW

Hierarchy evaluation process
Case Study BTH

- Streaming service vs. Messaging service
- Alternatives: WLAN; UMTS; GPRS

**Streaming Service**
- Many packets
- Sustained throughput important

**Messaging Service**
- Very few packets
- Link capacity of minor importance
- Fast delivery important

<table>
<thead>
<tr>
<th></th>
<th>WLAN</th>
<th>UMTS</th>
<th>GPRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>0.138</td>
<td>0.298</td>
<td>0.264</td>
</tr>
<tr>
<td>Cost</td>
<td>0.311</td>
<td>0.349</td>
<td>0.340</td>
</tr>
<tr>
<td>Accessibility</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
</tr>
<tr>
<td>Preference Ratio</td>
<td>0.361</td>
<td>0.327</td>
<td>0.312</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>WLAN</th>
<th>UMTS</th>
<th>GPRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>0.587</td>
<td>0.189</td>
<td>0.224</td>
</tr>
<tr>
<td>Cost</td>
<td>0.320</td>
<td>0.363</td>
<td>0.317</td>
</tr>
<tr>
<td>Accessibility</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
</tr>
<tr>
<td>Preference Ratio</td>
<td>0.413</td>
<td>0.295</td>
<td>0.291</td>
</tr>
</tbody>
</table>
IMS Interworking

IMS interworking

- Between 3GPP and WLAN
- Between 3GPP and UMTS
- Between 3GPP and CDMA2000

Main ideas

- Extend 3GPP services and functionality to other environments
- Develop bearer services allowing 3GPP subscribers to use other environments to access 3GPP PS services
3GPP UMTS/WLAN Architecture

Interworking architectures for 3GPP UMTS/WLAN

- Tight coupling
- Loose coupling
- P2P architecture
ROVER

- New architectural solution, called ROVER, suggested by BTH for L5 HO with mobility prediction

- ROVER: Routing in OVERlay networks

Goals

- Enable mobile users to seamlessly move among networks of diverse technologies, while maintaining the service continuity and the QoS across application and IP domains

- Provide support for both unicast and multicast services, with particular focus on content distribution purposes
ROVER (cont’d)

- Project initially supported by .SE (2007/2008), and now part of FP7 EU STREP PERIMETER (2008)

- Focus: media distribution in overlay networks (initially) and L5 HO (today)

- Particular focus
  - QoS-aware overlay routing
  - Middleware
  - Mechanisms for media distribution
  - Study of protocols for multicast distribution

- Blekinge Institute of Technology (BTH) team
  - Professor Adrian Popescu
  - TeknDr David Erman
  - TeknDr Doru Constantinescu (now with HiQ, Karlskrona)
  - TeknLic Dragos Ilie (now with Business Security, Lund)
  - PhD student Alex Popescu
  - MSc Karel de Vogeleer
ROVER Architecture
Research Challenges

- Middleware
- Overlay routing
- BitTorrent media distribution
- Overlay multicast networks
- Interworking platform
- Vertical handover
- Mobility modeling and prediction
- Decision-making algorithms
- Handover security
Several Important Results

- Partial implementation of a dedicated middleware
- Framework named Overlay Routing Protocol (ORP) suggested to provide a QoS-aware service on top of IP’s best effort service
- Simulation study of ORP
- Modifications and extensions suggested to the BitTorrent (BT) to make it suitable for use in providing a streaming video delivery service
- Simulation study of the suggested BT modifications and extensions
- Comparative simulation study of three representative categories of overlay multicast networks, i.e., Application Layer Multicast Infrastructure (ALMI), Narada and NICE
ROVER Middleware

- Middleware: software that bridges and abstracts underlying components of similar functionality and exposes this through a common API
- Object-oriented (C++) based API
- Based on the Key-Based Routing (KBR) of the common API framework suggested by the authors of CHORD
- Intended to work on top of both structured and unstructured underlays; compared to this, the initial KBR was suggested to work only on top of a structured underlay
- Quick integration of existing protocol implementations
- Development, evaluation, testing, performance analysis of different protocols and combinations of protocols
Unicast QoS Routing in Overlay Networks

- Particular difficulties
  - Multiple constraints
  - Dynamic environments; presence of churn
  - “Real-time” performance demand

- QoS constraints can be
  - Additive (e.g., for delay), or
  - Multiplicative (e.g., for packet loss), or
  - min-max (e.g., for bandwidth)

- Optimization algorithms
  - Self-Adaptive Multiple Constraints Routing Algorithms (SAMCRA)
  - The Simplex Method
  - Gradient Projection Method
  - Conjugate Gradient Method
  - Particle Swarm Optimization
**Unicast QoS Routing in Overlay Networks (cont’d)**

- **ROUTING**: the process of selecting paths in a network such that they satisfy a set of simultaneous QoS constraints
  - Routing algorithms: given a network topology, find the desired paths
  - Routing protocols: ensure that all nodes have “accurate” topology information

- Example: “Find a path from node A to node B with a minimum of 1Mbps capacity, such that the delay does not exceed 100ms and the packet loss probability is no higher than 0.01%”

- Types of path QoS metrics (example: path i→j→k→…..→l→m)
  - Additive (delay, jitter): \( d(i,j)+d(j,k)+…+d(l,m) \)
  - Multiplicative (packet loss): \( 1-(1-p(i,j))(1-p(j,k))…(1-p(l,m)) \)
  - Min-max (bandwidth): \( \min(c(i,j), c(j,k), \ldots, c(l,m)) \)
Unicast QoS Routing in Overlay Networks (cont’d)

- Research has been done on
  - Finding paths suitable for transporting multimedia flows
  - Path selection is done such as to satisfy a set of simultaneous QoS constraints
  - Reacting to path failures by reallocating flows to backup paths
  - Implementing this functionality in an overlay network spawned by end-nodes, without changing existing Internet infrastructures
Unicast QoS Routing in Overlay Networks (cont’d)

- Study has been done on
  - Flow allocation problems and optimization algorithms
  - Gnutella measurements and characteristics modeling
  - Overlay Routing Protocol (ORP) framework
    - Route Discovery Protocol (RDP): finds QoS-constrained paths by selective forwarding
    - Route Maintenance Protocol (RMP): handles churn by reallocating flows to backup paths
  - Different performance metrics have been evaluated, e.g.,
    - Call blocking ratio, bandwidth utilization, bandwidth overhead, path stretch (RDP)
    - Path failure ratio, restored paths ratio, bandwidth utilization, bandwidth overhead (RMP)

- The experiments have shown that RDP and RMP are viable alternative to provide a QoS-aware service at the application layer

- The cost has been observed to be maximum 1.5% of the residual network capacity

- Future work regards implementation of RDP and RMP and PlanetLab tests
Handover

- BTH has developed a solution for vertical handover, called Network Selection Box (NSB)
- NSB encapsulates the raw packet in UDP and sends it over a real network
- A tunnel is used to send the packets over the interfaces encapsulated in UDP
- NSB can be used for the transport over WLAN, UMTS and GPRS
Conclusions

- Very fascinating and complex research!
- Opening for many applications based on “telepresence”, e.g., pay-free system, check in-free system
- Need for move towards real-live deployment, e.g., PlanetLab
- Need for participation in the standardization efforts
References


Semi-Official Logo

Thanks to Eric Jacobson
THANK YOU!