Routing in Large-Scale Self-organized Networks

Serge Fdida
LIP6-CNRS / Université Paris 6
http://www.lip6.fr/rp/~sf

NeXtworking'03

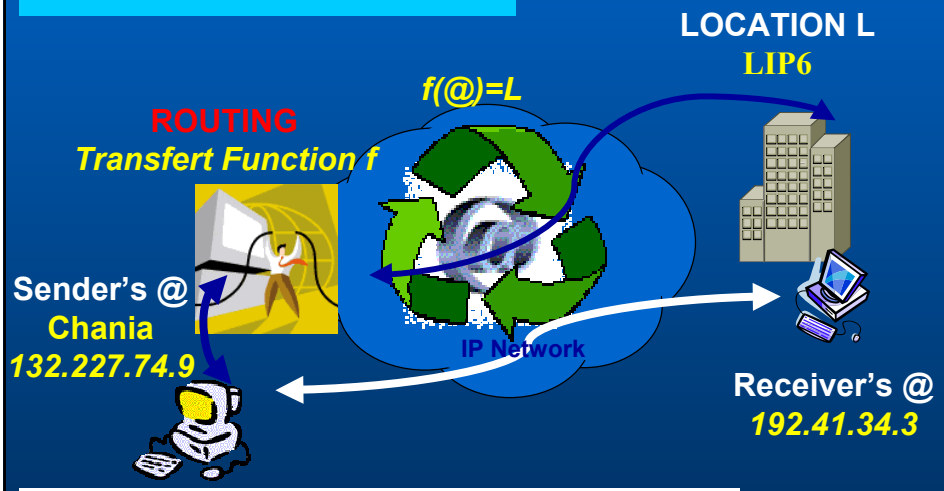
Future Internet Issues

- **Routing is hard!**
  - **BGP**
  - **Multicast**
  - **Mobile**
  - **Ad-Hoc**
  - **P2P**

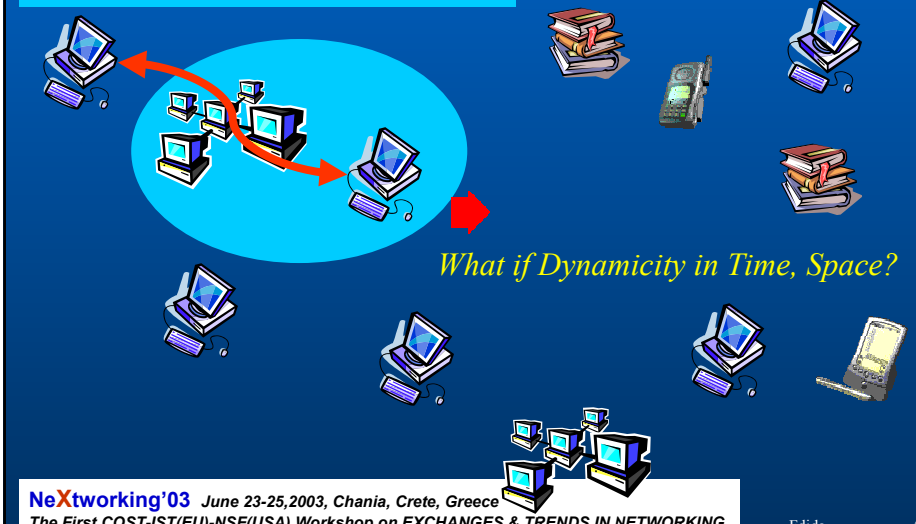
Routing formalization?

- **? We need to re-assess the relationship between**
  - **Address**
  - **Location (physical)**
  - **Route computation**
  - **State complexity**
- **? Decouple**
  - **Physical network topology**
  - **« Logical » network topology**

Routing semantics and operation

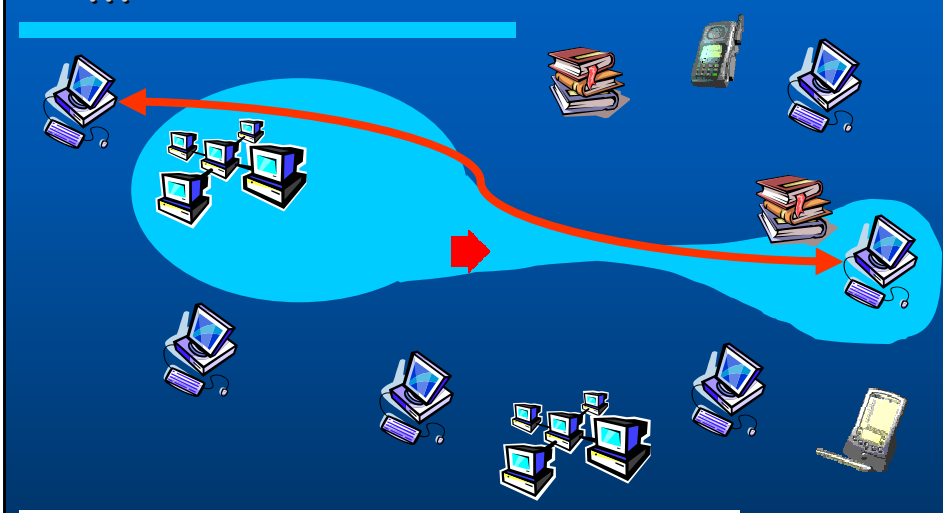


ROUTING = Transfer Function



What if Dynamicity in Time, Space?

Mobile, Multicast, QoS, Content Access, ...



Examples:

Node X, Network adress @_X, Location L

- **IP routing**
  - AS_x = f(IP@_X), f={RIP, OSPF, IS-IS}
  - If X Moves! f does not return « L »
- **Mobile IP**
  - Location L = f(g(IP@_X), g=HA, f=FA
  - Two transfert functions are required (convolution)
- **Geographic routing**
  - Location L = f(GPS_X), GPS_X =(Lat,Long)
  - GPS required, but Stateless

Exemples:

Node X, Network adress @_X, Location L

- **Route Server**
  - @_X = get(Serv), @Serv is known
  - L = f(@_X), Two-phase
  - MPOA/NHRP
- **Content access**
  - Data = f(key), f=Distributed Hash Table
  - f : See Chord, Pastry, Can, …

New routing paradigms

- **Very Large Scale**
  - Human Users & Machines
  - Ambiant networking, BGP scaling,
  - *Complexity in the routing table states*
- **Mobile**
  - Same reasons plus PDAs, Phones, embedded devices
  - *Convolution of transfert functions*
- **Overlays**
  - Simpler to deploy, Content access
  - *Dedicated*

Main problems

- **Addressing structure**
  - How many addresses per node
  - Of what type, use?
- **Impact on the topological space**
- **Routing in a mathematical space**
- **Mapping a multi-dimensional data to a one-dimensional value**
- **Robust to mobility**

Illustration with an example: Indirect routing using distributed location information

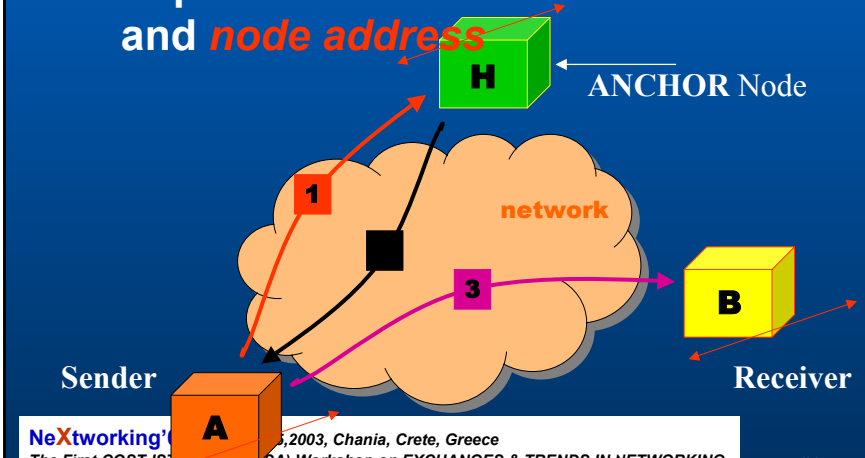
Aline Viana[1,2], Marcelo Amorim[1], Serge Fdida[1], and José Rezende[2]

LIP6 Laboratory
University of Paris VI
www.lip6.fr

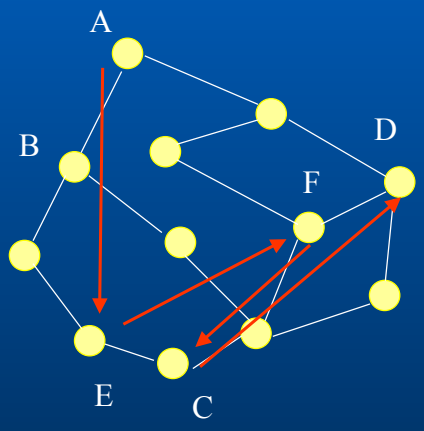
GTA/COPPE
Fed. Univ. Rio de Janeiro
www.gta.ufrj.br

Indirect routing

- **Separation between *node identifier* and *node address***



Physical / Virtual Topology



A →	E	
	F ↓	
	C →	D
X		

Node's roles

- **Identification**
  - **Universal identifier U**
    - Uniquely identifies a node in the real system
  - **Virtual identifier V**
    - Uniquely identifies a node in the virtual topology
  - **Relative (topology-dependent) address E**
    - Mapping of U in a value belonging to the relative addressing space
    - Utilization of a DHT, known by all nodes of the topology

Topology creation

- The nodes are identified by their relative addresses, which are based on their neighborhood (mobile nodes)
- When a node joins the network, it receives a control region from one of its neighbors
- The addressing space is a segment [0, 2^n[

$R_i^{\oplus} \subset V = [0, 2^m[$

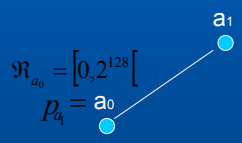
The relative address *E*: identifies the relative location of the node ($E_i = R_i^{\ominus}$)

For the first node at t=0: $[R_0^{\ominus}, R_0^{\oplus}[= [0, 2^m[$

Topology creation

- Node *n* first identifies its neighbors when it joins the network
- Among these neighbors, the one which has the largest region will become the *n*'s *parent neighbor*
- The parent neighbor then gives to *n* a part of its own control region

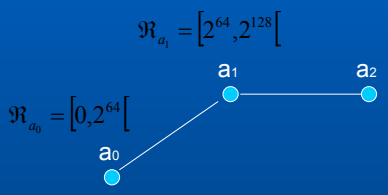
Topology creation (exemple)



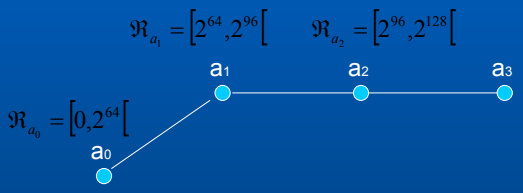
$A_{a_0} = \phi \qquad A_1 = \{a_0\}$

$L_{a_0} = \phi \qquad L_1 = \{ [E_{a_0}, \Re_{a_0}] \} = \{ 0, [0,2^m[\}$

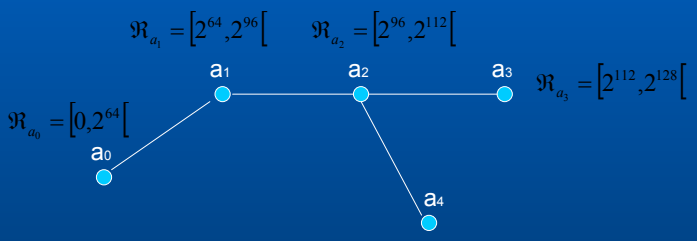
Topology creation (exemple)



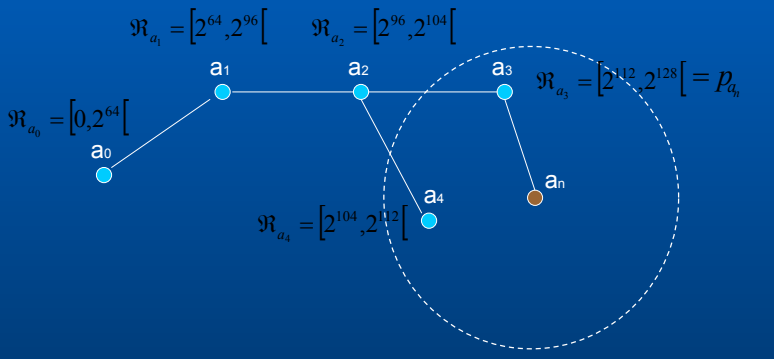
Topology creation (exemple)



Topology creation (exemple)



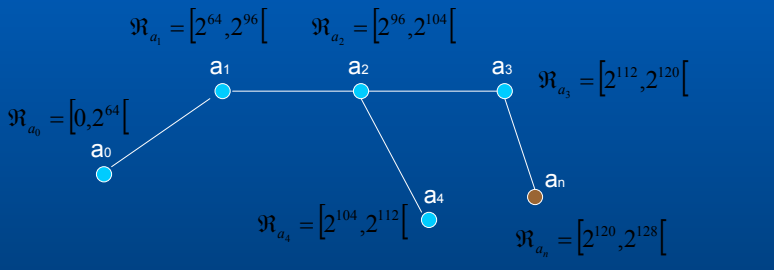
Topology creation (exemple)



$$A_n = \{a_3, a_4\}$$

$$L_n = \{ [2^{104}, [2^{104}, 2^{112}[\,], [2^{112}, [2^{112}, 2^{128}[\,] \}$$

Topology creation (exemple)



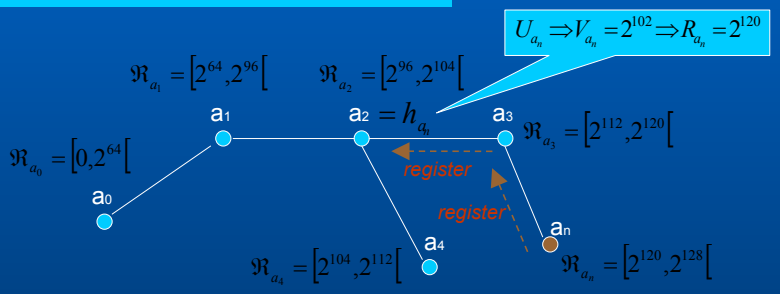
Types of address

- The universal identifier U
- The virtual identifier V
  - Used for identifying a node's *anchor*
  - The Anchor node behaves as a Home Agent for a set of nodes in its controlled region
- The relative address E
  - Identifies a unique node in the logical network
  - Changes when node moves
  - Used for routing

Locating nodes

- The *n*'s anchor is node *h* whose control region contains the virtual address *Vn*
  - Node h is identified by its relative address *Eh*
- Node *a* wants to contact node *b*
  - *a* knows $V_b=f(U_b)$
  - *a* sends a search message to the neighbor whose control region gets the message as close as possible to *Vb*
  - The message is routed hop by hop until it reaches the node that contains *Vb*
- The anchor node responds to *a* with a message containing the current *b*'s relative address
  - Node *b* had already informed *h* about its current position

Tribe register procedure

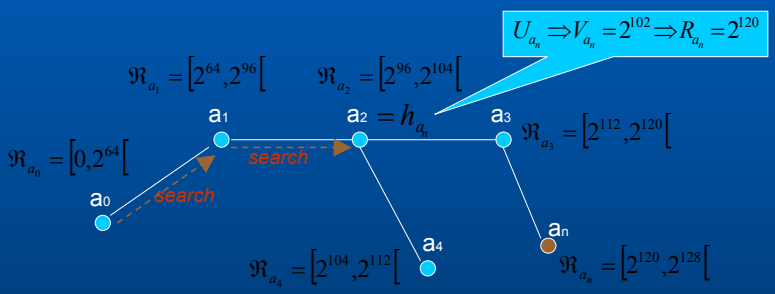


$$E_{a_n} = 2^{120}$$

$$f(U_{a_n}) = V_{a_n} = 2^{102}$$

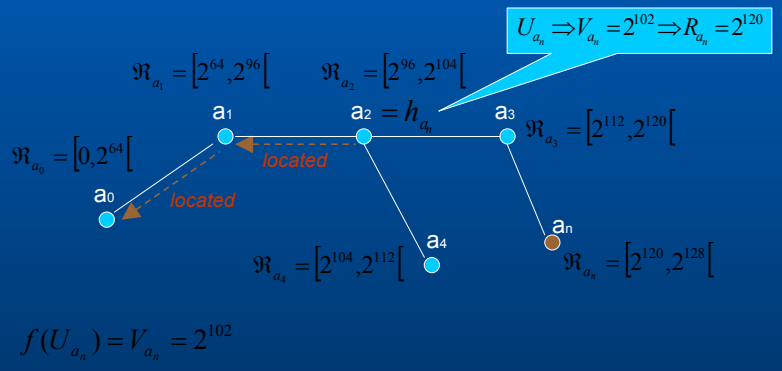
$$h_{a_n} = 2^{96}$$

Tribe location procedure



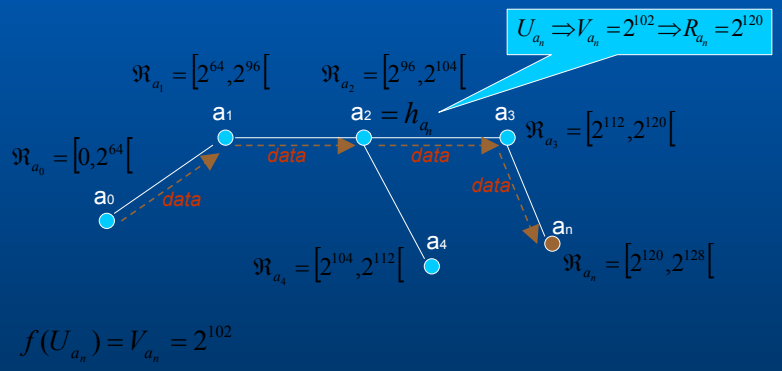
$$f(U_{a_n}) = V_{a_n} = 2^{102}$$

Tribe location procedure



$$f(U_{a_n}) = V_{a_n} = 2^{102}$$

Tribe location procedure

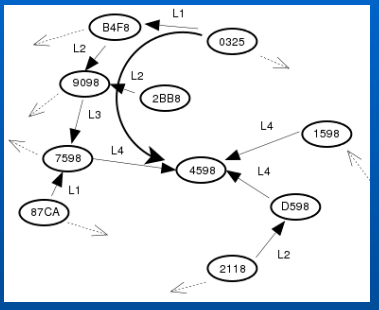


$$f(U_{a_n}) = V_{a_n} = 2^{102}$$

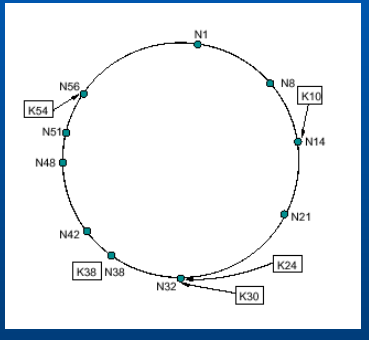
Region continuity

- How can you enforce region continuity when a node leaves/moves?
- Need a multi-dimension space
- Should be robust to mobility

Logical Topologies

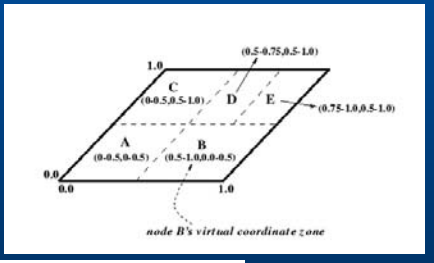


TAPESTRY

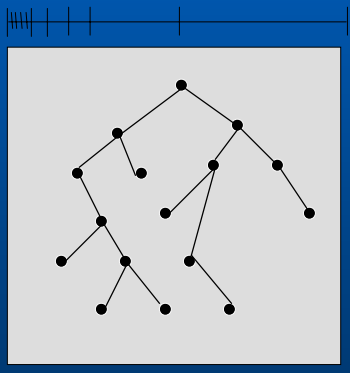


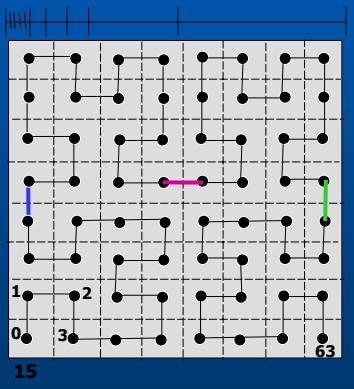
CHORD

CAN



The virtual space





Requirements

- Consecutively ordered points should be adjacent in space
- Can represent a space of multi-dimensions in one dimension
- The space should be partitioned in a recursive way,
  - A node leaving the network don't cause a inconsistency of the routing procedure
- Many paths can be used

Reference

- **Indirect Routing Using Distributed Location Information**
  Viana Aline c., Dias de amorim Marcelo, Fdida Serge and Rezende José F.
  IEEE International Conference on Pervasive Computing and Communications (PerCom)
  Dallas-Fort Worth, Texas - March, 2003

- **http://www-rp.lip6.fr**
  - **publications**