A New Routing Specification for Packet Radio Datagram Networks

Andreas Gal (DB7KG) [1]

TheNetNode-Group [1], NORD><LINK e.V., Braunschweig

Abstract

Though many different implementations of datagram-based packet routers exist, nearly all of them are based on the NET/ROM protocol specification. However, with growing network size and transport capacity the original concept has reached its limit. To acknowledge the new environment, a new time-based routing protocol has been designed to meet the needs of today's ham radio packet network. The *Internode Protocol* described here is a full compatible modern replacement for NET/ROM, with a fast network wide information propagation combined with low bandwidth requirements. The *Internode Protocol* is intended to be no proprietary solution but a well-documented standard.

Introduction

The *Internode Protocol (INP)* is a „drop-in“ compatible protocol to NET/ROM. It is possible to operate in the same network with mixed routing software, nevertheless maximum stability and performance is reached in homogeneous networks.

Information exchange

Neighbor communication

All information traffic between two network neighbors is performed by AX.25 frames with PID[1] 0xCF (INP is concerned to be still a NET/ROM-class protocol). The detection of new capabilities is done with protocol elements not defined in NET/ROM (i.e. L3RTT-frames), thus no collisions with existing implementations are expected. All information is exchanged with numbered frames in connected state (I-type frames), each layer 3 frame is transported piggy-back on an AX.25 frame sent from one neighbor to another. This connection between the layer 3 neighbors is called *interlink*.

This *interlink* transports normal layer 3 information frames (NET/ROM syntax) and additional *INP* information.

The layer 4 communication is done with NET/ROM standard.

Link, Link Quality

The quality of an interlink is determined by its mean frame transport time, measured by return-timer frames (RTT) divided by two. This value is called *smoothed neighbor transport time (SNTT)*. The frames are sent by both neighbors[2]. The interlink measure is done with fixed time interval. Each measured value is smoothed with the last values, new established links should start with higher values to prevent race conditions. On temporal link interruptions, the last value can be stored as start value for the smoothing procedure. The neighbor must respond within 180 seconds to an RTT-query or the link is re-established (AX.25 link reset).

Routes, Definition and Interpretation

Routes are the most meaningful part of a network. Network nodes are presenting each other their known destinations by sending route information about these targets. This is called *route time (RT)*. The sum of this value and the smoothed neighbor transport time to the propagating neighbor represent the *target time (TT)*. The best path from node A to B is determined by node A by searching the neighbor with the lowermost target time.

Routes can have a target time from 10ms up to 599,99s with a granularity of 10ms. 600s is considered the horizon of routing information. No target times above 600s are sent (*target time limit*).

Routes with target time below the horizon are called positive information and represent available nodes in the network. A routing information is valid as long as it is updated by more recent information through the originated neighbor. The loss of the interlink connection to a neighbor or the reception of a target time of 600s marks a route as unusable.

---

[1] Protocol Identifier

[2] The frame format is described later in this document.

Any update of a positive information with slower target time (or route loss) is called negative information.

A route time above 600s is never sent, but should be treated as negative information.

Because the own target time depends on the smoothed neighbor transport, positive and negative information can originate from RTT-measurements, too.

The effective target time limit can be reduced locally, any target time above this value is treated and sent like the horizon value (600s)[3].

Hop-Counter, Definition and Interpretation

On an imaginary path through the network each transporting router is called a *hop*. The nodes are counting and propagating this *hop-counter*, to give a second information horizon. When an information reaches a configured hop-counter value, it is treated to be negative[4].

Routing Concept

The *Internode Protocol* defines only triggered information updates, thus information is valid as long as the originated interlink connection is usable or it is updated. To prevent deadlock situations, each node must be able to predict the behavior of its neighbors in some situations. The postulations written here must not be violated.

Net of Trust

The receiver of information accepts any information told by the sender with minimum validity checks. He *trusts* the sender.

Responsibility of the Sender

The sender of information is *responsible thereof*. Positive information has to be negotiated to the receiver, if the own route is lost. This responsibility is recursive for each interlink relation.

Segmentation of the Network

The neighbors of a node can be considered as windows to the network. There are forward and backward routes for each of these windows.Responsible for forward routes is the node, for backward routes the neighbor.

Redundancy reduction, Poison Reverse

A backward information is never returned to the neighbor. The node can send the best (maybe slower) alternate route (alternate reverse) or propagate the destination as not available (poison reverse).

Alternate reverse is the best choice for fast networks with stable interlinks. Poison reverse should be used in slow or busy networks.

The algorithms can be mixed in a network.

Penalty

Each propagation of a route implicates the increase of the route time by one granularity step (10ms). The hop-counter is increased by one.

Predictable target time

The target time is the sum of the smoothed neighbor transport times of each transporting router (hop).

Propagation scheduling

Negative information is propagated with higher priority then positive information. The propagation has to occur immediately after arrival. To reduce information volume, the propagated value can be increased by an adaptive value to catch possible following decrements (*negative preload*).

Positive information can be delayed, as it is not time critical. Updates should be sent only if the transport time improved by a high perceptual value and the difference is above a minimum absolute level[5]. The bandwidth used for positive information should be adopted to the receiving neighbor's interlink quality. For interlinks with low bandwidth, the amount of data can be reduced by sending only fast routes with low target time. This reduces the responsibility of the sending node, because not sent information has not to be reverted.

Negative information have to be propagated inside an acceptable delay (*guaranteed maximum delay of negative information*, i.e. 10s). If this is not possible (i.e. the connection to the receiver is busy), the interlink connection must be re-established[6].

Predictable information range

Routing information is propagated only as long as their hop-counter does not exceed a locally

---

[3] This makes it possible to reduce the routing table size.

[4] Some protocols (TCP/IP RIP) use only hop-counters for routing decision.

[5] It makes no sense to propagate an improvement of 50% from 20ms to 10ms. TheNetNode i.e. is propagating route improvements of 25%, with a minimum absolute decrease of 100ms.

[6] This situation should be prevented by adoptive positive information propagation (reduction of the information volume).

bytes	**1**	**7**	**1**	**2**	**n**	**1**
data	**0xFF**	AX.25 shifted	1 .. 255	1 .. 60000, MSB first		**0x00**
description	signature	callsign	hop counter	transport time	<opt. Fields>	**EOP**
	RIF	RIP (inside RIF)				

Figure 1: Format of a RIF. Several RIPs can follow (only one showed here).

bytes	**1**	**1**	**n**
description	field length	field type	field data

Figure 2: Format of the optional fields. Several of these can be included in a RIP.

field name	**length**	**type**	**contents**
Alias	variable	0x00	node alias (without tailing spaces), ASCII
IP	5	0x01	IP address (network format) and byte packed hostmask

Figure 3: Defined optional fields. More optional fields will be released later.

configured value (*hop-counter limit*). This hop-counter limit has to be adjusted to the expected route lengths in the network[7]. A positive route with exceeding hop-counter is propagated like negative routes.

Routing Decision

Every node is sending Layer 4 information addressed to a known node to the neighbor with the lowermost target time. This is called the best route, or the **primary route**.

Route deletion

Routes can be deleted, if both, forward and backward information, are negative.

Node deletion

Nodes can be deleted, if there are neither forward nor backward positive route information for this node[8].

Routes exchange

Information about routes are exchanged by *routing information packet* (RIP). Several of these frames can be gathered to a single AX.25 Frame, called *routing information frame* (RIF).

Routing Information Frame

A RIF is a numbered information frame on the interlink connection, starting with a single identification byte of 0xFF (signature). This value is guaranteed not to appear in normal L3 frame headers as the first byte. The RIF is filled with RIPs, partial RIPs or segmentation of a RIP is not allowed (**fig. 1**).

Routing Information Packet

A RIP concludes mandatory and optional fields. Figure 1 shows a RIF including a RIP.

Each optional fields is identified by the overall option length and option type. Options can be skipped by the receiver without knowledge of the option format.

The shortest possible RIF contains the callsign of the propagated route, the hop-counter, the target time, and the end-of-packet (**EOP**) field (**fig. 2**).

For the own node, a RIP is created and sent after the interlink has been established. The hop-counter has to be 1, the target time is ignored by the receiver.

Optional Fields

Optional fields are used to exchange additional information on a node[9].

Optional fields received with unknown type are ignored, but must be stored for propagation. The number of optional fields in a RIP is limited by the RIF size only.

---

[7] TheNetNode has a limit of 30 hops.

[8] Unreachable nodes can be stored for layer 7 information purposes.

[9] The node Alias is a good example for additional information.

bytes	7	7	1	4	n	1
data	AX.25 shifted	AX.25 shifted	0x02	0x00 0x00 0x00 0x00	Text	0x10
description	source node call	L3RTT	opcode	dummy header data	ASCII	CR

Figure 4: L3RTT frame format.

flag description	syntax
INP capable	$N
IP v*X* frames are accepted on the interlink	$I*X*

Figure 5: Defined neighbor flags. Additional flags will be released later.

The Callsign field is filled with the node callsign in shifted AX.25 syntax. The transport time is sent divided by 10ms with the most significant byte first (**fig. 3**).

Optional fields are sent only together with positive route information, and only the optional information of the primary route is accepted. This is called the *primary optional fields*.

Optional Fields Propagation

Only primary optional fields are propagated, changing these fields (by a new source neighbor or fields change them) implicates a optional fields update to all neighbors.

Optional fields are propagated with the first positive information to a network segment, further route updates are sent without optional fields, to reduce propagation bandwidth.

Neighbor RTT measurement

Layer 3 RTT measurement frames are transported inside normal layer 3 information frames. The layer 3 destination address is „L3RTT-0“, to prevent collision with previous implementations (**fig. 4**).

Inside the data area, a text field is transported. This field contains implementation specific information (for RTT measurement) and the INP *flags*. The text field is filled up with spaces (ASCII 0x20) to reach MTU[10].

The neighbor will reflect a received RTT frame immediately and unchanged. The runtime of each RTT frame divided by two is the neighbor transport time.

[10] Maximum Transmission Unit

The text field of the RTT frame, received from the neighbor, is scanned for flags only. Each flag starts with „$“ character, followed by the flag type and additional flag data (**fig. 5**).

Any INP node must send the flag type „**N**“ to mark itself INP capable. This mechanism is used to detect INP neighbors.

Interlink phases

Interlink Establishing

Two neighbors not having an active interlink connection have only negative forward and backward information to each other.

Both neighbors try simultaneously to establish the interlink connection[11]. On success, interlink established phase is entered.

Interlink Established

Once an interlink connection has been established successfully, each neighbor sends immediately a RTT measurement frame.

As long as the RTT measurement has not been successful, no further data is sent or accepted on the interlink.

AX.25 link resets are ignored during the establishing phase.

Interlink Failure

On a AX.25 link failure, all routes to and from this neighbor are dropped. The neighbor detection mechanism should be restarted. Interlink establishing phase is entered.

Interlink Reset

[11] NET/ROM nodes broadcast can be used to detect neighbors.

On a AX.25 link reset, all routes are dropped and the interlink established phase is entered.

Acknowledgment

This protocol was created with the support of Hans Georg Giese (DF2AU), Joachim Scherer (DL1GJI), Peter Stirnimann (HB9PAE), and many others.

The RTT-measurement algorithm was created by Peter Gülzow (DB2OS) and others for the TheNetNode project and adopted for INP.

The RIF/RIP format was developed by Joachim Scherer (DL1GJI).

Thanks to Alexander Kurpiers (DL8AAU) for language fine-tune.

Literature

[1] Hans Georg Giese (DF2AU),
NET/ROM protocol specification, appendix F of the TheNetNode user manual, printed 1997.

[2] Andreas Gal (DB7KG),
Critical review of the poison reverse algorithm used by XNET, published in AMPR, 1997

[3] Andreas Gal (DB7KG),
Roadmap TheNetNode 1.70, published in AMPR by Peter Gülzow (DB2OS), 1997

[4] Nils Kassube (DF6LN),
New NET/ROM broadcast format, published in AMPR, 1995

[5] Hans Georg Giese (DF2AU),
RFC for the 3rd developer meeting in Kassel, published in AMPR, 1997

[6] Henning Happe (DG9FU),
TheNetNode user manual, printed version April 1997.

[7] Joachim Scherer (DL1GJI),
INP routing information frame format, published in AMPR, 1997