Parametric Analysis of the PGM Protocol

Petr Matoušek, Mihaela Sighireanu

`matousp@fit.vutbr.cz, sighirea@liafa.jussieu.fr`

Brno University of Technology, Czech republic
LIAFA, Paris University 7, France

1. Introduction

❖ **Talk outline**

1. **Protocol PGM**
2. **Modeling PGM**
3. **Parametric Analysis**
4. **Verification**
5. **Conclusion**

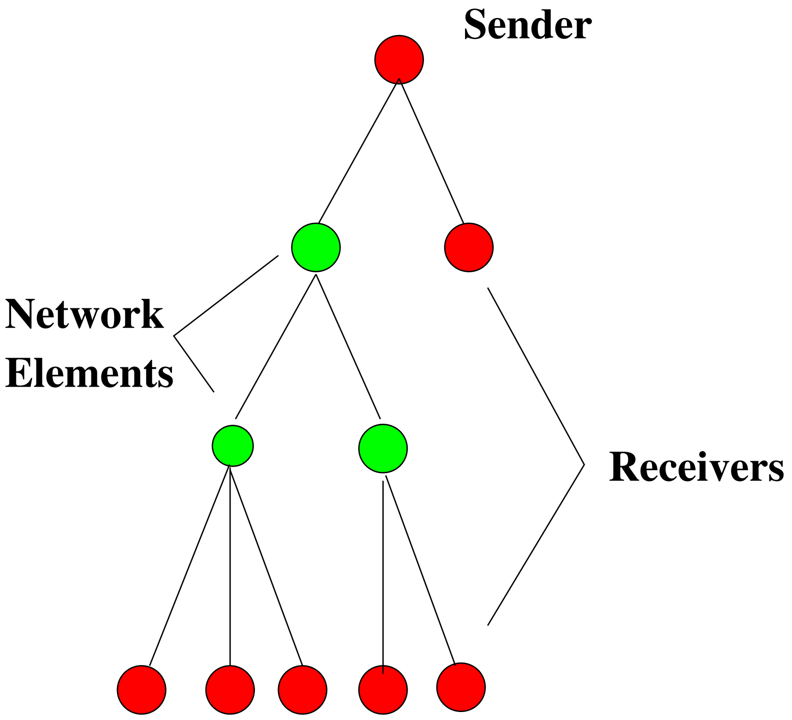
1. Protocol PGM

❖ General Overview

- **PGM (Pragmatic General Multicast) defined by RFC 3208.**
- **Reliable multicast transport protocol for application, that require ordered or unordered, duplicate-free, multicast data delivery from multiple sources to multiple receivers .**
- **Members may join and leave the group at any time.**
- **Many different types of data packets (ODATA, RDATA, SPMs, NAKs etc).**

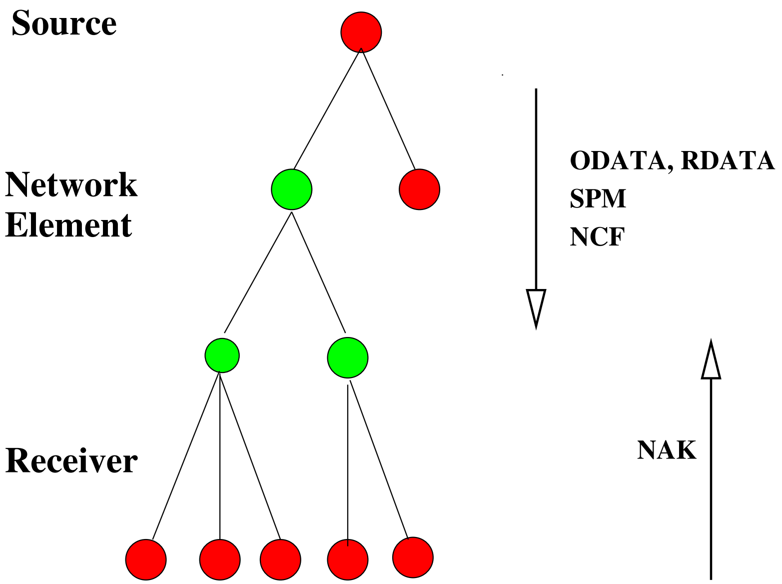
1. Protocol PGM - Introduction

- **Protocol Architecture**



1. Protocol PGM - Introduction

❖ Data Transmission



- Data (`ODATA`, `RDATA`),
- `SPM` (Source Path Message),
- `NAK` (Repair request), `NCF` (`NAK` confirmation)

2. Protocol PGM - Verification

❖ **PGM guarantees that "a receiver either receives all data packets from transmissions and repairs, or is able to detect unrecoverable data packet loss".**

❖ **Several verification studies on PGM has been done.**

❖ **B.Bérard, P.Bouyer, and A. Petit: Analysis the PGM protocol with** UPPAAL. **RT-TOOLS, August 2002.**

- **Verification of a simplified timed version of PGM with linear topology and one-placed buffer.**
- **The reliability property of the protocol is verified by instantiating the parameters and calling the** UPPAAL **tool.**
- **Verification of two properties:**

2. Protocol PGM - Verification [BBP]

❖ **Lost info property** - “For each data, each receiver knows if it did receive the data or if it will never receive it”.

- 960 control states, 5 clocks, 25 bounded variables
- Property `E<>(obs.Error)` is True means the receiver may make mistake to estimate restoration of a data.

❖ **No-loss property** - “Each data which is detected as lost is eventually repaired”.

- 17280 control states, 5 clocks, 35 bounded variables
- Property `E<>(receiver1.test==1 ro receiver2.test==2)` is True, that means it is not verified.

1.Protocol PGM - Verification [BS]

- M.Boyer, M.Sighireanu: Synthesis and verification of constraints in the PGM protocol. FME, September 2003. (ADVANCE, 2nd year)
  - Verification of the PGM using classical tools (IF, CADP).
  - Manual synthesis of the constraints between parameters.
  - Verification of full reliability property using TREX .
  - Property verified by instantiation of parameters.
  - Analysis of complexity - addressing of sources of complexity.
- Our goal: To obtain the constraint deduced in this work automatically.

1.Protocol PGM - Verification [BS]

❖ **Losses-signaled property** - "a receiver either receives all data packets and repairs, or is able to detect unrecoverable data packet loss".

- The property was verified for all messages, except for those of the last transmission window - a problem of closing window.
- The problem can be solved using "closing `SPM` ".

❖ **Parametric analysis of full reliability property** - finding a relation between parameters of the system that satisfies the property.

- The relation (a constraint with parameters) was manually derived.
- The property was successfully verified using instantiation of the parameters - the result confirmed the property.

1.Protocol PGM - Verification [others]

❖ P.Boigelot, L.Latour: Verifying PGM with infinitely many packets. LIAFA 2002.

- Validation using LASH of the sliding window mechanism of the protocol for any number of data packets sent.
- Different model based on finite state automate - no time model.
- Study the relation between the LEAD and TRAIL values of the Transmit Window and Receive Window.

❖ J.Esparza, M.Maidl: Simple representative instantiations for multicast protocols. TACAS, 2003.

- Mathematical framework for multicast protocol that allows to generalize the results obtained for linear topologies to tree topologies.

2. Modeling PGM

❖ **Analysing the full PGM protocol is beyond limits of current verification tools because of**

- **dynamic topology - joining/leaving a node,**
- **multiple senders,**
- **a lot of different packet types (`SPMs`, `NCF` , `NAKs`),**
- **a lot of processes, counters and clocks.**

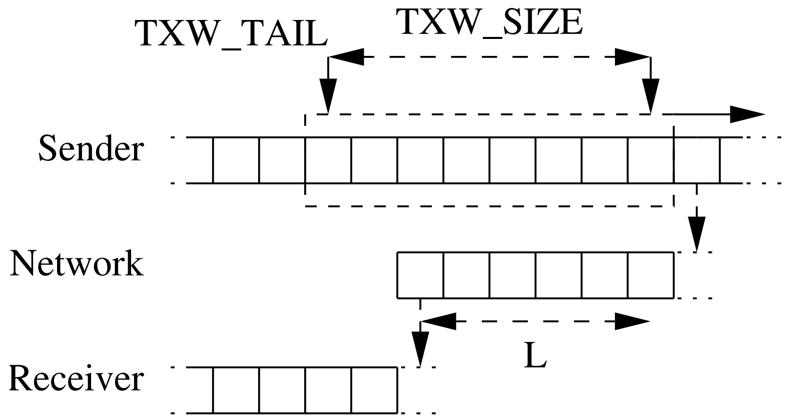
❖ **Sources of complexity:**

- **many variables,**
- **non-linear constraints.**

❖ **We need a new abstract model.**

2. Modeling PGM - An abstract model

❖ **The abstract model is based on a global view of the protocol running between the sender and one of the receivers**



- **Linear topology - a sender, network, a receiver.**
- **Network is abstracted into unreliable, unbounded FIFO queue implemented by a counter automaton.**
- **Only data packets (`ODATA`) are transmitted.**

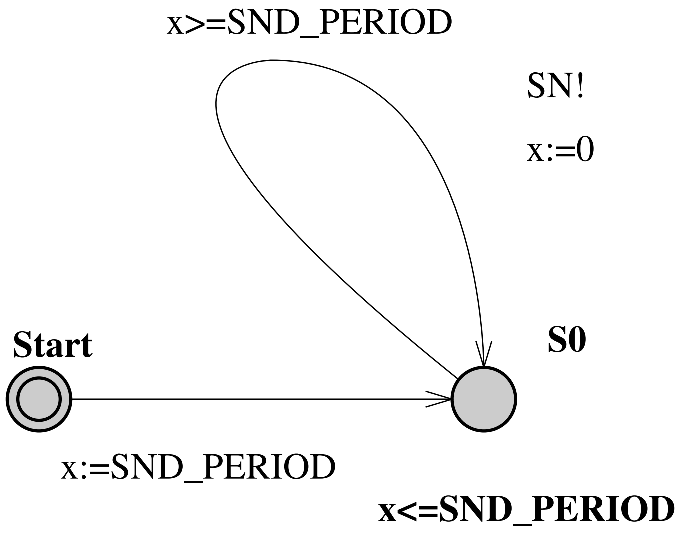
2.Modeling PGM - The abstract model

- **Global view abstraction reduces number of counters and variables.**
- **Clocks, counters, variables:**
  - **two clocks -** x, y**, two counters -** L, def_lost**,**
  - **one finite variables -** lp**,**
  - **six parameters - `RATE`, `NLOSS`, `TXW_SIZE`, `BUFFER_LENGTH`, `SND_PERIOD`, `CH_PERIOD`.**

2. Modeling PGM - The sender

❖ The sender

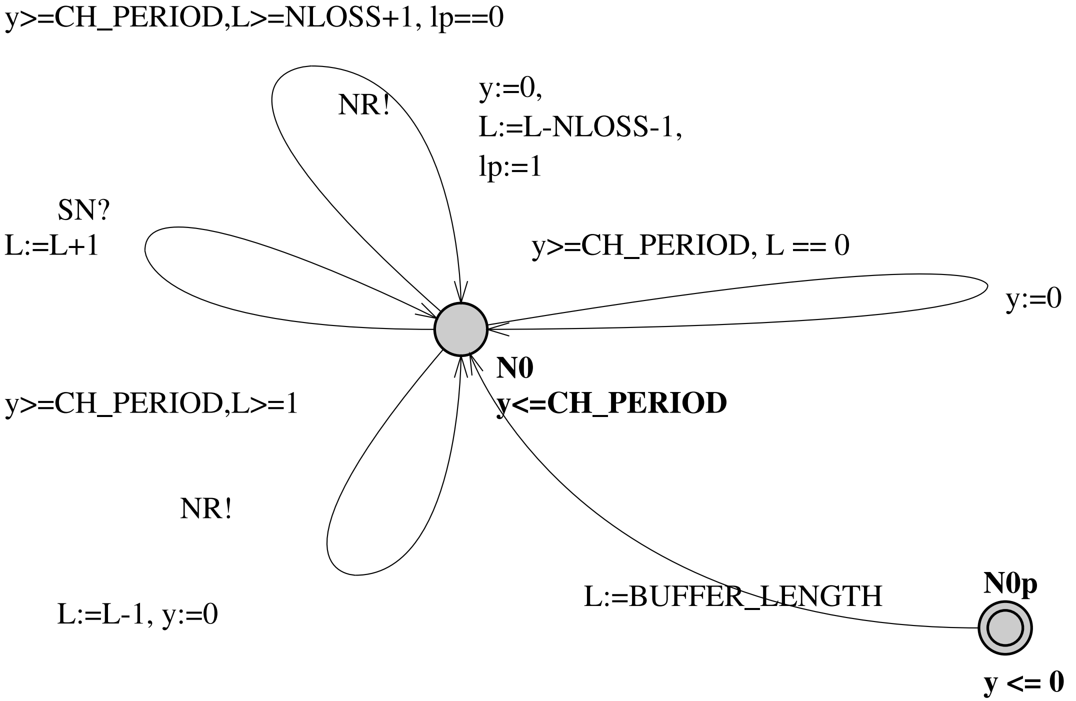
- **generates ODATA packets each** `SND_PERIOD`**,**
- **advances trasmitting window by one after each data packet is sent.**
- **The transmit window is fixed in order to save data as long as possible.**



2. Modeling PGM - The network

❖ The network

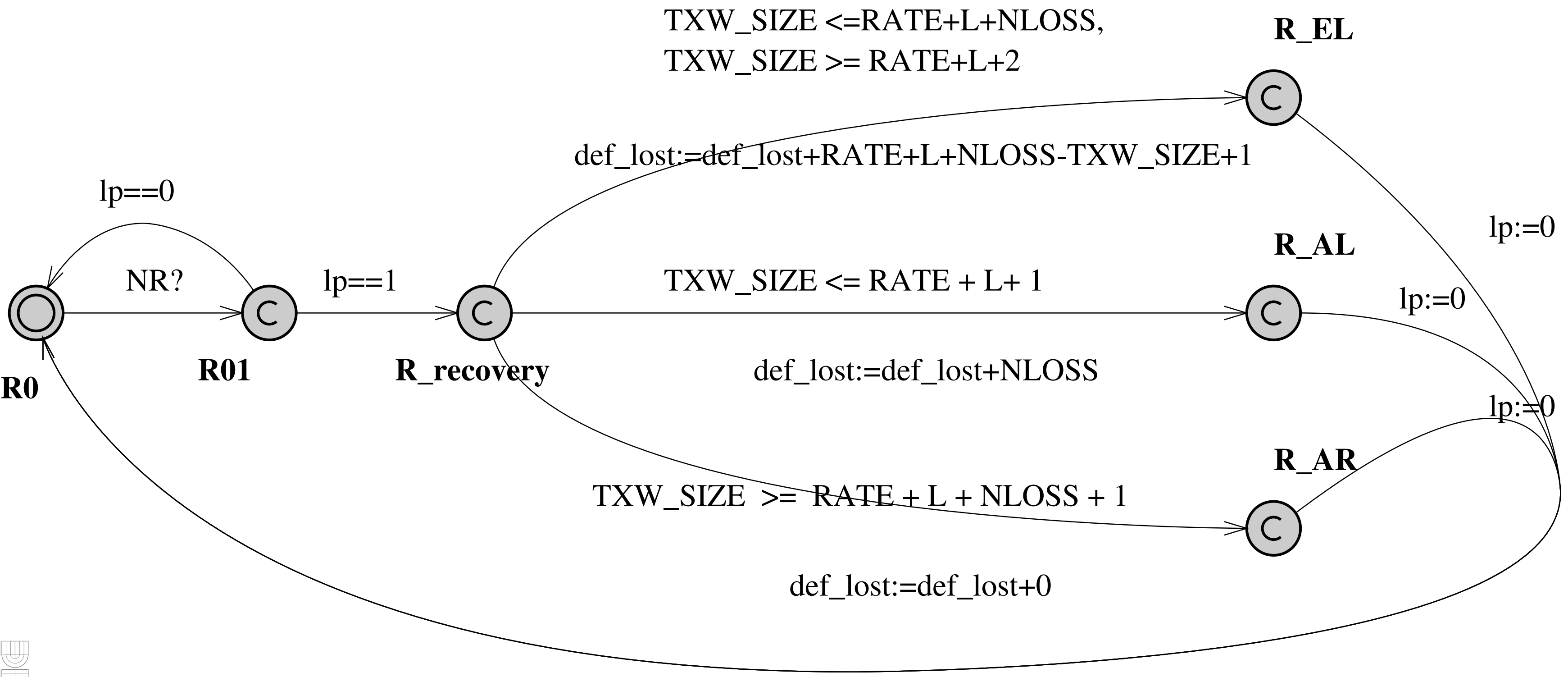
- **receives data from the sender,**
- **delivers data to the receiver each** `CH_PERIOD`**,**
- **non-deterministically generates losses of NLOSS packets (variable lp)**



2. Modeling PGM - The receiver

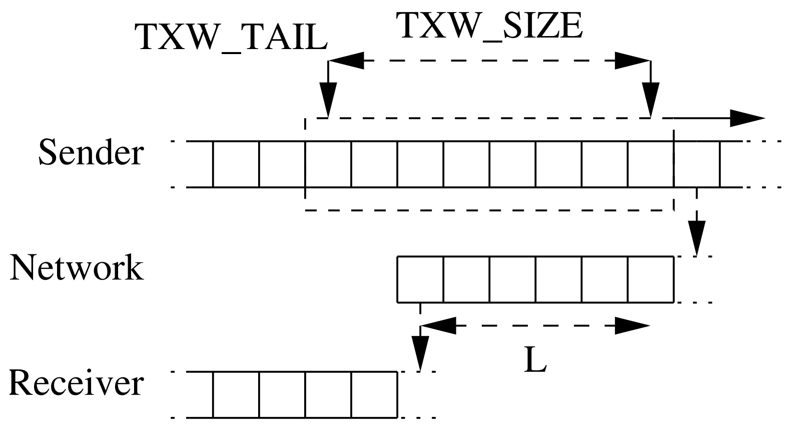
❖ **The receiver**

- **accepts data from the network,**
- **detects losses - computes if lost packets can be recovered.**
- **`RATE` is ratio between the transmission speed and `SND_PERIOD`.**



2. Modeling PGM - Detection of losses

❖ Global view abstraction



∀R All lost packets may be recovered if

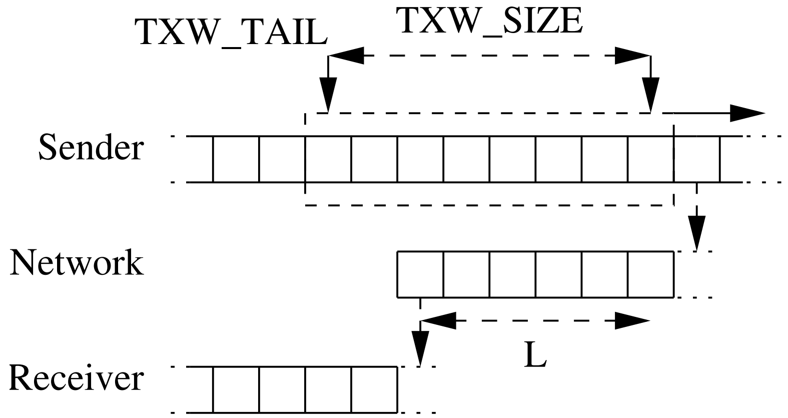
$$\texttt{TXW\_SIZE} > \texttt{RATE} + L' + \texttt{NLOSS}$$

∀L None of the `NLOSS` lost packets may be recovered if

$$\texttt{TXW\_SIZE} \leq \texttt{RATE} + L' + 1$$

2. Modeling PGM - Detection of losses

- **Global view abstraction**



∃R Some of the lost packets may be recovered if

$$\begin{aligned}\texttt{TXW\_SIZE} &> \texttt{RATE} + L' + 1\\ \texttt{TXW\_SIZE} &\leq \texttt{RATE} + L' + \texttt{NLOSS}\end{aligned}$$

- **Only first relation satisfies the full reliability property.**

3. Parametric Analysis

- All lost packet may be recovered if

$$\mathtt{TXW\_SIZE} \quad > \quad \mathtt{RATE} + L' + \mathtt{NLOSS}$$

where L' (the current value of L) is a variable, where $L' = L - \mathtt{NLOSS} - 1$.

⇒ This constraint must be satisfied by the parameters in order to obtain full reliability.

- But L is a variable - we need a relation depending only on time and parameters.
- L can be computed as follows

$$L \quad = \quad f(t, \mathtt{BUFFER\_LENGTH}, \mathtt{SND\_PERIOD}, \mathtt{CH\_PERIOD}, \mathtt{NLOSS})$$

3. Parametric Analysis

- To compute L, we distinguish four cases:

Case 1 $\texttt{SND\_PERIOD} > \texttt{CH\_PERIOD}$

- The rate of arrivals is less than departures.
- The size of the queue converges to zero by time.

$$0 \leq \quad L \quad \leq \texttt{BUFFER\_LENGTH}$$

Case 2 $\texttt{SND\_PERIOD} = \texttt{CH\_PERIOD}$

- Arrivals are the same speed as departures.
- The size of the queue decreases to a value less then `NLOSS` because of losses.

$$0 \leq \quad L \quad \leq \texttt{BUFFER\_LENGTH}$$

3. Parametric Analysis

Case 3 $\texttt{CH\_PERIOD}/\texttt{SND\_PERIOD} > \texttt{NLOSS}$

- Arrivals are faster than the sum of departures and losses.
- The queue grows beyond any limits by time.

$$\texttt{BUFFER\_LENGTH} \leq \quad L \quad < \infty$$

Case 4 $\texttt{NLOSS} > \texttt{CH\_PERIOD}/\texttt{SND\_PERIOD} > 1$

- Arrivals are faster than departures, but not enough to fill the losses between two delivery.
- The queue is alternating depending on non-deterministic losses.

$$0 \leq \quad L \quad < \infty$$

3. Parametric Analysis - Constraints

- **After substitution of L' and using limits on L we get following constraints:**
- **The constraint for full recovery is**

$$\mathtt{SND\_PERIOD} \geq \mathtt{CH\_PERIOD} \ \wedge \ \mathtt{TXW\_SIZE} \geq \mathtt{RATE} + \mathtt{BUFFER\_LENGTH}$$

- **Partial recovery of losses is possible if**

$$\mathtt{TXW\_SIZE} > \mathtt{RATE} + \mathtt{BUFFER\_LENGTH} - \mathtt{NLOSS}$$

- **None of losses may be recovered if**

$$\mathtt{TXW\_SIZE} \leq \mathtt{RATE} + \mathtt{BUFFER\_LENGTH} - \mathtt{NLOSS}$$

3. Parametric Analysis - Conclusion

- The constraints between parameters and the law of evolution of L are non-linear relations on reals and integers.
  - For instance, exact value of L for case 3 is

$$L = \mathtt{BUFFER\_LENGTH} + \left[\frac{t}{\mathtt{CH\_PERIOD}} \right] \left[\frac{\mathtt{CH\_PERIOD}}{\mathtt{SND\_PERIOD}} - 1 - \mathtt{NLOSS} \right]$$

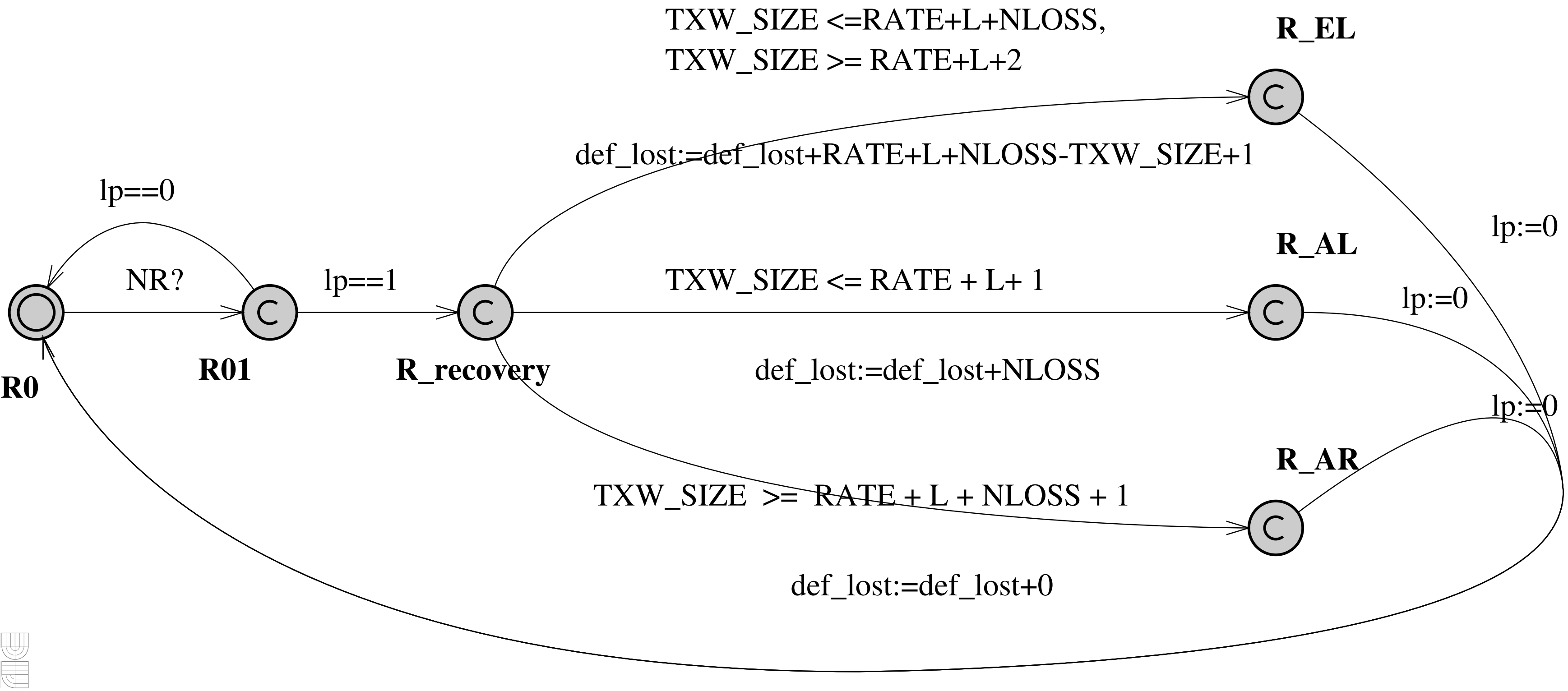
- Verification can be done
  - by instantiating some of parameters to avoid non-linear constraints,
  - by applying acceleration,
  - by applying approximation.

4. Verification - One Time Loss Model

- **Full reliability property** - **“a receiver either receives all data packets or it is able to recover all lost data packets.”**
- **“One Time Loss” Model**
  - **Modified model where a loss appeared just once per session.**
  - **Non-linearities reduced using instantiation of some parameters.**
  - **To speed up analysis we carefully set initial conditions.**
- **For parametric verification of the model we use HYTECH and TREX**

4. Verification - One Time Loss Model

- **Three extended TA communicating via synchronization - the sender, the network, the modified receiver.**
- **One finite variable (lp), two clocks (x,y), two counters (def_lost, L).**
- **Six parameters.**



4. Verification - using HyTech

❖ Parametric verification using HyTech

- HyTech is a tool for parametric verification of hybrid systems.
- HyTech does not support acceleration - generation of full reachability set does not terminate.
- To test our property we need to define a final region
  `final_reg := def_lost > 0`
  where the property is violated.
- We can get only results where the property is not satisfied.

❖ HyTech output (for partial losses, CH_PERIOD/SND_PERIOD>= 2)

RATE>= 1 & SND_PERIOD>= 1 & BUFFER_LENGTH>= 1 &
CH_PERIOD<= 2 SND_PERIOD&
TXW_SIZE+ NLOSS>= RATE+ BUFFER_LENGTH+ 3 &
SND_PERIOD< CH_PERIOD& NLOSS<= BUFFER_LENGTH+ 1
& TXW_SIZE<= RATE+ BUFFER_LENGTH+ 2

4. Verification - using TREX

- **TREX is a tool for parametric verification of timed systems.**
- **Model is based on extended timed automata.**
- **TREX generates a set of reachable configuration for the input model and finite symbolic graph.**
- **It uses efficient extrapolation techniques to accelerate computation:**

$$\begin{aligned} C &= \{2 \leq x \leq 6, 1 \leq y \leq 4\} \\ post_\theta(C) &= \{2 \leq x \leq 6, 1 \leq y \leq 6\} \\ post_\theta^2(C) &= \{2 \leq x \leq 6, 1 \leq y \leq 8\} \\ post_\theta^*(C) &= \{2 \leq x \leq 6, 1 \leq y \leq 4 + 2 * n\} \; using \; periodicity \end{aligned}$$

- **Data structure in TREX are represented using Parametric DBMs (PDBMs).**

4. Verification - using TREX

- Case 1: `SND_PERIOD` $>$ `CH_PERIOD`
  - **R_AR**
    txw_size $\geq$ rate + buffer_length
    and buffer_length $\geq$ nloss + 1
  - **R_EL**
    txw_size $\geq$ rate + buffer_length -nloss - n3 - 1 and
    twx_size $\leq$ rate + buffer_length - n3 - 3 and
    buffer_length $\geq$ nloss + n3 - 3 and
    buffer_length $\geq$ n3 - 2 and
    n3 $\geq$ 0
  - **R_AL**
    txw_size $\leq$ rate -nloss + buffer_length - n3 - 1 and
    buffer_length $\geq$ nloss + n3 - 2 and
    buffer_length - n3 - 1 $\leq$ 0 and
    n3 $\geq$ 0

4. Verification - using TREX

- **Case 2: `SND_PERIOD` = `CH_PERIOD`**
  - **R_AR**
    txw_size ≥ rate + buffer_length
  - **R_EL**
    txw_size ≥ rate + buffer_length -nloss + 1 and
    twx_size ≤ rate + buffer_length - 1
  - **R_AL**
    txw_size ≤ rate - nloss + buffer_length

- **No acceleration needed in this case.**

4. Verification - using TREX

- **Case 3: $\texttt{CH\_PERIOD}/\texttt{SND\_PERIOD} > \texttt{NLOSS}$**
  - **New parameter $q = \texttt{CH\_PERIOD}/\texttt{SND\_PERIOD}$, we consider $q \geq 2$**
  - **Constraints similar like in the first case.**
- **Case 4: $\texttt{NLOSS} > \texttt{CH\_PERIOD}/\texttt{SND\_PERIOD} > 1$**
  - **The experiments results are similar to the third case.**

4. Verification - Conclusion

- We successfully verified One Loss Time Model
- Analysis of the Full Abstract Model
  - There is no way to always recover losses in case 3 and case 4.
  - This can be done by searching a graph of symbolic configurations where $def\_lost = 0$.
  - The problem is to generate this graph - L is complex, so the automatic computation fails.
- Another interesting point - the number of definitively lost packets
  - To compute that number we need a class of assignments for counters - not possible for DBMs.
- We need another data structure !

5. Conclusion

- **Future direction - parametrized intervals**
  - **Based on Interval Diagrams extended with parameters.**
  - **Domain is a vector (like PDBMs) with pair of bounds.**
  - **New abstract data structure - p-hcubes**
    - **used for representation of configurations on counters (PDBMs for clocks)**
    - **space representation in $O(n)$ - better than PDBMs $O(n^2)$**
    - **canonical representation**

- **It will be a part of a new version of TReX.**

5. Conclusion

❖ Parametric verification of PGM protocol

1. New abstract model of PGM protocol based on global view of the system.
2. Parametric analysis of the system
    - Synthesis of constraints on parameters that satisfies the full reliability property.
    - Detection of non-linear relations between parameters ⇒ instantiation.
    - “One time loss” model.
3. Full automatic verification of the model with parameters using TREX and HYTECH .
4. To verify Full Abstract Model we need a new data structure - we propose parametrized intervals.