# Proceedings of the Design, Analysis, and Simulation of Distributed Systems Symposium, April 2-8, 2005, San Diego, USA, c. 72-78. Verification of Protocol BGP via Decomposition of Petri Net Model into Functional Subnets

Dmitry A. Zaitsev

Odessa National Telecommunication Academy Kuznechnaya, 1, Odessa 65029 Ukraine http://www.geocities.com/zsoftua

# ABSTRACT

Petri net model of the widely known protocol BGP of Internet backbone routing was constructed. The decomposition of Petri net model of communication protocol BGP into functional subnets was implemented. Invariance of the source model was proved on the base of established invariance of its functional subnets. The speed-up of computations obtained is exponential with respect to dimension of Petri net.

**KEYWORDS:** Petri net, BGP, Invariance, Decomposition, Functional subnet

# **1. INTRODUCTION**

Petri nets [1] are used successfully for the investigation of distributed systems and concurrent processes in different application fields [2,3]. Detailed models of reallife objects have huge enough dimension, numbering hundreds of elements. At that time, basic methods of Petri net properties analysis have an exponential calculation complexity. It makes the analysis of real-life systems difficult.

Models of complex systems are assembled of models of its components usually. Moreover, in the cases the composition of model out of subnets is not given, we suggest to apply methods of Petri net decomposition represented in [4]. The algorithm of decomposition allows the partition of a given Petri net into set of its functional subnets. Our decomposition differs from known approaches [5,6,7] in the class of components named by functional subnets.

The technique the invariants of functional subnets, defining partition of source Petri net, are applied for calculation of the entire Petri net invariants was studied in [8]. The speed-up of computations obtained is estimated with exponential function. Since the dimension of subnets, as a rule, is essentially lesser than the dimension of the entire net, the actual speed-up of computation may be extremely considerable that was confirmed by the results of this technique application to communication protocols analysis [9,10].

The present paper constitutes in essence a case study of decomposition technique via verification of the widely known protocol BGP [11] of Internet backbone routing. Petri net model of BGP protocol constructed is simplified enough but it allows the accurate implementation of the decomposition technique. Moreover, as distinct from [9,10], non-minimal functional subnets are used.

In Section 2 the brief description of BDG protocol is represented; in Section 3 the Petri net model of protocol is constructed; in Section 3 the decomposition of model into functional subnes is realized; in Section 5 the decomposition-based calculation of place invariants of model is implemented; in Section 6 invariants of transitions are calculated on the base of decomposition of dual Petri net; in Section 7 the speed-up of computations is estimated and in Section 8 conclusions are formulated.

### 2. COMMUNICATION PROTOCOL BGP

The Border Gateway Protocol (BGP) [11] is an interautonomous system routing protocol. It is the very significant for the whole Internet operability, so the autonomous systems constitute a backbone of the global data exchange. More than thirty RFC (Requests For Comments) are devoted to BGP protocol specification and refinement. Recently the most widespread is BGP-4 [12], but the distinctions in comparison with the first standard specification [11] are the very specific and inessential for a draft model construction.

The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. This network reachability information includes information on the autonomous systems (AS's) that traffic must transit to reach these networks. This information is sufficient to construct a graph of AS connectivity from which routing loops may be pruned and policy decisions at an AS level may be enforced.

There are five types of standard BGP messages:

- 1 OPEN,
- 2-UPDATE,
- 3 NOTIFICATION,
- 4 KEEPALIVE,
- 5 OPEN CONFIRM.

After a transport protocol connection is established, the first message sent by either side is an OPEN message. If the OPEN message is acceptable, an OPEN CONFIRM message confirming the OPEN is sent back. Once the OPEN is confirmed, UPDATE, KEEPALIVE, and NO-TIFICATION messages may be exchanged.

UPDATE messages are used to transfer routing information between BGP peers. The information in the UP- DATE packet can be used to construct a graph describing the relationships of the various autonomous systems. By applying rules to be discussed, routing information loops and some other anomalies may be detected and removed from the inter-AS routing.

BGP does not use any transport protocol based keepalive mechanism to determine if peers are reachable. Instead KEEPALIVE messages are exchanged between peers often enough as not to cause the hold time (as advertised in the BGP header) to expire. The KEEPALIVE message is a BGP header without any data.

NOTIFICATION messages are sent when an error condition is detected.

### **3. MODEL OF PROTOCOL BGP**

Petri net model of protocol BGP is represented in fig. 1. Let's remind, that *Petri net* [1] is a triple N = (P,T,F), where  $P = \{p_1,...,p_m\}$  – finite set of nodes named places,  $T = \{t_1,...,t_n\}$  – finite set of nodes named transitions, flow relation  $F \subseteq P \times T \cup T \times P$  defines a set of arcs connecting places and transitions. Thus, Petri net is directed bipartite graph; one part of nodes consists of places, another – of transitions. Places are drawn as circles, transitions – as bars. Usually, graph N is supplemented with a marking defining an initial arrangement of tokens in places. Tokens are dynamic elements that move inside net as a result of transitions firing.

In the general case a *net with multiply arcs* is considered. It contains an additional mapping  $W : F \to N$ . Multiplicity, in a case it is distinct from unit, is pointed as a number *w* on the corresponding arc. Flow relation and arcs' multiplicities may be represented via incidence matrix. Let us introduce matrices  $A^-$ ,  $A^+$  of input and output arcs of transitions accordingly:

$$\begin{split} A^{-} &= \left\| a^{-}_{i,j} \right\|, \ i = \overline{1,m}, \ j = \overline{1,n}; \\ a^{-}_{i,j} &= \begin{cases} w(p_i, t_j), \ (p_i, t_j) \in F, \\ 0, \ otherwise. \end{cases} \\ A^{+} &= \left\| a^{+}_{i,j} \right\|, \ i = \overline{1,m}, \ j = \overline{1,n}; \\ a^{+}_{i,j} &= \begin{cases} w(t_j, p_i), \ (t_j, p_i) \in F, \\ 0, \ otherwise. \end{cases} \end{split}$$

And finally, we introduce an *incidence matrix* A of Petri net as  $A = A^+ - A^-$ .

Special notations of sets of input and output nodes for places and transitions are introduced also:

• 
$$p = \{t \mid \exists (t, p) \in F\}, p^{\bullet} = \{t \mid \exists (p, t) \in F\},\$$
  
•  $t = \{p \mid \exists \{p, t\} \in F\}, t^{\bullet} = \{p \mid \exists (t, p) \in F\}.$ 



Figure. 1. Petri net model of protocol BGP

The model describes asymmetric interaction of two systems. First system is represented with places  $p_1 - p_5$ and transitions  $t_1 - t_6$ , second system – with places  $p_6 - p_{10}$  and transitions  $t_7 - t_{12}$ . Places  $p_{11} - p_{14}$  correspond to communication subsystem and model standard messages: OPEN, OPENCONFIRM, and KEEPALIVE. Notice that the model represents only procedures of connection establishment and maintenance, abstracting of data transfer for adjustment of routing tables. Date interchange is implemented in state ESTABLISHED with the aid of standard messages UPDATE. This process is not displayed in model constructed. Semantic description of elements of the model is represented in Table 1.

### 4. DECOMPOSITION OF MODEL

The decomposition of model into functional subnets is represented in fig. 2. Notice that four drawn functional subnets  $Z^1$ ,  $Z^2$ ,  $Z^3$ ,  $Z^4$ , defining a partition of source model, are not minimal. As the result of algorithm described in [4] application we obtain the decomposition into minimal subnets induced by the subsets  $\{t_1\},\{t_2,t_5,t_6\},\{t_3\},\{t_4\},\{t_7\},\{t_8,t_{11},t_{12}\},\{t_9\},\{t_{10}\}$ . So, for instance, subnet  $Z^2$  constitutes a sum of two minimal subnets induced by transitions  $t_3$  and  $t_4$  correspondingly. Problems of the functional subnets composition out of the minimal functional subnets were studied in [4].

Let's remind, that the functional net [4,8] is a special case of net with input and output places. Functional net is a triple Z = (N, X, Y), where N – is Petri net,  $X \subseteq P$  – input places,  $Y \subseteq P$  – output places, besides sets of input and output places do not intersect:  $X \cap Y = \emptyset$ , moreover, input places do not have input arcs, and output places do

not have output arcs:  $\forall p \in X : {}^{\bullet}p = \emptyset$ ,  $\forall p \in Y : p^{\bullet} = \emptyset$ . Places of a set  $C = X \cup Y$  are named by *contact*, and places of a set  $Q = P \setminus (X \cup Y)$  are named by *internal*.

Table 1. Description of model's elements

Element	Description
$p_1, p_6$	Initial state of systems
<i>p</i> <sub>2</sub>	Open request sent
<i>p</i> <sub>7</sub>	Open request received
$p_3, p_8$	Connection established
$p_4$	KEEPALIVE message sent
<i>p</i> <sub>9</sub>	KEEPALIVE message received
<i>p</i> <sub>5</sub>	KEEPALIVE message received
$p_{10}$	KEEPALIVE message received
<i>p</i> <sub>11</sub>	OPEN message
<i>p</i> <sub>12</sub>	OPENCONFIRM message
$p_{13}, p_{14}$	KEEPALIVE message
t <sub>1</sub>	Send OPEN message
<i>t</i> <sub>7</sub>	Receive OPEN message
t <sub>8</sub>	Send OPENCONFIRM message
<i>t</i> <sub>2</sub>	Receive OPENCONFIRM message
$t_3, t_{10}$	Send KEEPALIVE message
$t_4, t_9$	Receive KEEPALIVE message
$t_5, t_{11}$	Connection keep alive loop
$t_6, t_{12}$	Disconnection

Functional net Z = (N', X, Y) is named a *functional* subnet of net N and is denoted as  $Z \succ N$  if N' is subnet of N, and, moreover, Z is connected with the residuary part of the net only through the arcs incident with either input or output places, besides input places may have only input arcs and output places – only output arcs. Thus

$$\forall p \in X : \{(p,t) \mid t \in T \setminus T'\} = \emptyset,$$
  
 
$$\forall p \in Y : \{(t,p) \mid t \in T \setminus T'\} = \emptyset,$$
  
 
$$\forall \in Q : \{(p,t) \mid t \in T \setminus T'\} = \emptyset \land \{(t,p) \mid t \in T \setminus T'\} = \emptyset.$$

Functional subnet is named a *minimal*, if it does not contain any other functional subnet. According to theorem 2 proved in [4], any functional subnet Z' of a Petri net N is a sum (union) of a finite number of minimal functional subnets. Thus, a set of minimal functional subnets is the generating family for a set of functional subnets of a given Petri net N.

Subnet Z = B(R) = (X, Q, Y, R) of Petri net N is a *complete* in N, iff in N the following conditions hold true:  $X^{\bullet} \subseteq R, \ ^{\bullet}Y \subseteq R, \ ^{\bullet}Q^{\bullet} \subseteq R$ .

#### Algorithm of decomposition:

Step 0. Choose an arbitrary transition  $t \in T$  of the net N and include it in the set of selected transitions  $R := \{t\}$ .

Step 1. Construct subnet Z generated by the set R: Z = B(R) = (X, Q, Y, R).

Step 2. If Z is the complete in N, then Z is sought subnet. Stop.

Step 3. Construct the set of absorbed transitions:

$$S = \{t \mid t \in X^{\bullet} \land t \notin R \lor t \in Y \land t \notin R \lor t \in Q^{\bullet} \land t \notin R\}$$

Step 4. Assign  $R := R \bigcup S$  and go to Step 1.

As it was proven in [4]:

- subnet Z is complete in Petri net N iff it is functional subnet of N;
- subnet Z constructed by the algorithm of decomposition is minimal functional subnet of Petri net N;
- the algorithm of decomposition has polynomial complexity  $o(n^3)$ , where *n* is the number of nodes of the net.

# **5. INVARIANCE OF MODEL**

Invariants [1] are a powerful tool of the structural properties of Petri nets analysis. It allows the determination of boundness, safeness of net, necessary conditions of liveness and absence of deadlocks. These properties are significant for real-life objects' analysis, especially, for communication protocols [1-3].

*p-invariant* of Petri net [1] is a nonnegative integer solution of the system

$$\overline{x} \cdot A = 0 . \tag{1}$$

*t-invariant* of Petri net is a nonnegative integer solution of the system

$$\overline{y} \cdot A^T = 0$$
.

As according to [1] each t-invariant of Petri net is pinvariant of dual net, so further, not limiting the generality, we shall consider p-invariants only.

All the known methods of invariants calculation [13,14] have an exponential complexity. It makes the application of these methods to real-life objects' models, numbering thousands of elements, analysis difficult.

According to theorem 2 proved in [8], Petri net N is invariant iff all its minimal functional subnets are invariant and, moreover, exists a common nonzero invariant of contact places. Therefore, to calculate invariants of a Petri net it is required to calculate invariants of its minimal functional subnets and then to find common invariants of contact places. It was shown, that results are true for an arbitrary set of functional subnets, defining a partition of the set of transitions of Petri net.



Figure. 2. Decomposition of BGP protocol model

Let a general solution for invariant of functional subnet  $Z^{j}$  has the form

$$\bar{x}^j = \bar{z}^j \cdot G^j, \tag{2}$$

where  $\overline{z}^{j}$  is an arbitrary vector of nonnegative integer numbers, and  $G^{j}$  is a matrix of basis solutions. Then, since each contact place is incident with not more than two subnets [8], the system of equations for calculation of common invariants of contact places has the form

$$\left\{ \overline{z}^i \cdot G^i_p - \overline{z}^j \cdot G^j_p = 0, \ p \in C \right\},$$
(3)

where *i.j* is the numbers of functional subnets, incidental to a place  $p \in C$ , and  $G_p^j$  is a column of matrix  $G_p^j$ , that corresponds to place p.

Therefore, variables  $\overline{z}^{j}$  become not free ones. Note that system (3) has the same form as the source system (1). Thus, it may be solved with above-mentioned methods. Suppose that  $\overline{z} = \overline{y} \cdot R$ , where *R* is a matrix of basis solutions of system (3), and  $\overline{y}$  is an arbitrary vector of nonnegative integer numbers. So, the general solution of system (1) according to (2) may be represented as

$$\overline{x} = \overline{y} \cdot H , \ H = R \cdot G . \tag{4}$$

With the aid of tool Tina [15] we obtain the following basis invariants of the subnets enumerated in fig. 2:

$$Z^{1}: (x_{1}, x_{2}, x_{3}, x_{5}, x_{11}, x_{12}) = (z_{1}^{1}, z_{2}^{1}) \cdot G^{1},$$

$$G^{1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix},$$

$$Z^{2}: (x_{3}, x_{4}, x_{5}, x_{13}, x_{14}) = (z_{1}^{2}, z_{2}^{2}, z_{3}^{2}) \cdot G^{2},$$

$$G^{2} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix},$$

$$Z^{3}: (x_{6}, x_{7}, x_{8}, x_{10}, x_{11}, x_{12}) = (z_{1}^{3}, z_{2}^{3}) \cdot G^{3},$$

$$G^{3} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

$$Z^{4}: (x_{8}, x_{9}, x_{10}, x_{13}, x_{14}) = (z_{1}^{4}, z_{2}^{4}, z_{3}^{4}, z_{4}^{4}) \cdot G^{4},$$

$$G^{4} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

The composition of the model is defined by fusion of eight contact places indicated in fig. 2. Let us construct the system of equations for contact places:

$$\begin{array}{l} p_3: \ z_1^1+z_2^1-z_1^2-z_3^2=0, \\ p_5: \ z_1^1+z_2^1-z_1^2-z_2^2=0, \\ p_8: \ z_1^3-z_1^4-z_2^4=0, \\ p_{10}: \ z_1^3-z_1^4-z_3^4=0, \\ p_{11}: \ z_2^1-z_2^3=0, \\ p_{12}: \ z_2^1-z_2^3=0, \\ p_{13}: \ z_3^2-z_3^4-z_4^4=0, \\ p_{14}: \ z_2^2-z_2^4-z_4^4=0. \end{array}$$

The basis solutions of the system with respect to vector  $(z_1^1, z_2^1, z_1^2, z_2^2, z_3^2, z_1^3, z_2^3, z_1^4, z_2^4, z_3^4, z_4^4)$  have the form

Let us assemble the united matrix G of matrixes  $G^1$ ,  $G^2$ ,  $G^3$ ,  $G^4$ . Notice that matrix G may be constructed in different ways depending on the order of calculation of invariants for contact places. As each contact place is incident to two subnets, so its invariant may be calculated by two different ways.

1	1	1	1	0	1	0	0	0	0	0	0	0	0	0)	
	1	0	1	0	1	0	0	0	0	0	1	1	0	0	
	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	0	0	0	0	0	0	0	0	0	0	0	0	1	0	
<i>G</i> =	0	0	0	0	0	1	1	1	0	1	0	0	0	0	
	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
	0	0	0	0	0	0	0	0	1	0	0	0	0	0)	

After multiplication of matrixes we obtain:

 Notice that the source system has five basis solutions so sixth solution is the sum of second and fourth, and seventh – the sum of second and fifth.

Therefore, the model of BGP protocol is p-invariant so, for instance, invariant,

which is the sum of second, third and fourth basis invariants, contains all the natural components. Consequently, the model of protocol is safe and bounded. For all the states holds true  $\bar{x}^* \cdot \bar{\mu} = 3$ .

### 6. INVARIANTS OF TRANSITIONS

To calculate invariants of transitions we construct the dual Petri net (Fig. 3), decompose it (Fig. 4) and implement the technique described for place invariants. The decomposition contains six minimal functional subnets. For calculation of invariants, it is convenient to consider the decomposition into two functional subnets. Since subnet  $Z'^1$  consists of 9 transitions, we may compose remained minimal subnets into one subnet with 5 transitions.



Figure. 3. Dual Petri net of BGP protocol model

The following matrix represents the basis invariants of transitions:



Figure. 4. Decomposition of dual Petri net into minimal functional subnets

As, for instance, the sum of two basis invariants

contains all the natural components, so the model of protocol BGP is t-invariant. Therefore, the model is consistent. Sequence  $\sigma^* = t_1 t_7 t_8 t_2 t_3 t_9 t_{10} t_{11} t_4 t_5 t_3 t_9 t_{10} t_{12} t_4 t_6$ , corresponding to invariant  $\overline{y}^*$ , provides  $\overline{\mu}_0 \xrightarrow{\sigma^*} \overline{\mu}_0$ .

Notice that, though the model of protocol BGP is invariant, it contains deadlocks  $(p_2, p_8, p_{11})$  and  $(p_4, p_6, p_{13})$ , reached via sequences  $t_1t_7t_8t_2t_3t_9t_{10}t_4t_{11}t_6t_1$  and  $t_1t_7t_8t_2t_3t_9t_{10}t_4t_{12}t_5t_3$  correspondingly. It may be easily explained by the model does not represent timeouts provided by the source specifications. Supplied with transitions returning each system from the ESTABLISHED to the IDLE state the model becomes live.

#### 7. SPEED-UP OF COMPUTATIONS

Let us estimate the speed-up of computations obtained in the assumption of the exponential complexity of the algorithms [11,12] for the solving of linear Diophantine systems in nonnegative integer numbers. Let the complexity is about  $2^q$ , where q is the number of nodes of net.

Notice that even such rather tiny model allows the speed-up of computations. At calculation of place invariants, instead to solve the system of dimension 12, we solved five systems with the dimension not exceeding 8. If we not take into accounting polynomial multipliers, then we obtain sixteen fold  $(2^{12}/2^8 = 16)$  speed-up of computations.

Notice that, speed-up have been obtained for the net numbering about dozen of nodes. At investigation of large-scale nets, the speed-up may be rather huge [9,10], so it is estimated [8] as exponential function  $2^{m-r}$ , where  $r = \max_{i}(m_i, c)$  and  $m_i$  is the number of places of subnet

 $Z^i$ , c is the number of contact places.

### 8. CONCLUSION

Therefore, the decomposition of Petri net model of the communication protocol BGP into functional subnets was implemented. To verify the protocol, Petri net invariants were used. Invariance of the source model is proved on the base of established invariance of its functional subnets. Essential speed-up of computations obtained confirms the practical value of proposed technique.

# REFERENCES

- [1] Murata T. 1989. "Petri Nets: Properties, Analysis and Applications". *Proceedings of the IEEE*, April, vol. 77, 541-580.
- [2] Girault C.; Volk R. 2003. Petri nets for systems engineering A guide to modelling, verification and applications. Springer-Verlag.
- [3] Cortadella J.; Kishinevsky M.; Kondratyev A.; Lavagno L.; Yakovlev A. 2002. Logic synthesis of asynchronous controllers and interfaces. Springer-Verlag.
- [4] Zaitsev D.A. 2004. "Decomposition of Petri nets". *Cybernetics and Systems Analysis*, 43(5), 131-140.
- [5] Juan E.Y.T., Tsai J.J.P., Murata T. 1998. "Compositional verification of concurrent systems using Petri net based condensation rules". ACM Trans. on Programming Languages and Systems, 20(5), 917-979.
- [6] Christensen S., Petrucci L. 2000. "Modular analysis of Petri nets". *The Computer Journal*, 43(3), 224-242.
- [7] Haddad S., Ilie J.-M., Klai K. 2002. "An incremental verification technique using decomposition of Petri nets". In *Proceedings of the second IEEE International Conference on Systems, Man and Cybernetics*, October 6-9, Hammamet, Tunisia, vol. 2.
- [8] Zaitsev D.A. 2004. "Decomposition-based calculation of Petri net invariants". In Proceedings of Workshop on Token based computing of the 25-th International conference on application and theory of Petri nets, Bologna, Italy, June 21-25, 79-83.
- [9] Zaitsev D.A. 2004. "Verification of Protocol ECMA with Decomposition of Petri Net Model". In Proceedings of The International Conference on Cybernetics and Information Technologies, Systems and Applications, Orlando, Florida, USA, July 21-25.
- [10] Zaitsev D.A. 2004. "Verification of protocol TCP via decomposition of Petri net model into functional subnets". In Proceedings of the Poster session of 12th Annual Meeting of the IEEE / ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, October 5-7, Volendam, Netherlands, p. 73-75.
- [11] Loogheed K., Rekhter Y., Watson T.J. 1989. A Border Gateway Protocol (BGP). *RFC 1105*.

- [12] Rekhter Y., Watson T.J., Li T. 1995. A Border Gateway Protocol 4 (BGP-4). *RFC 1771*.
- [13] Colom J.M., Silva M. 1990. "Improving the linearly based characterization of P/T nets". In LNCS 483, ATPN.
- [14] Contejan E., Ajili F. 1997. "Avoiding slack variables in solving linear Diophantine equations and inequations". *Theoretical Computer Science*, 173, 183-208.
- [15] Berthomieu B., Ribet O.-P., Vernadat F. 2004. "The tool TINA – construction of abstract state space for Petri nets and Time Petri nets". *International Journal of Production Research*. (http://www.laas.fr/tina).