

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年7月2日(02.07.2020)



(10) 国際公開番号

WO 2020/137320 A1

(51) 国際特許分類:
H04L 12/70 (2013.01)

(21) 国際出願番号: PCT/JP2019/046289

(22) 国際出願日: 2019年11月27日(27.11.2019)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2018-243363 2018年12月26日(26.12.2018) JP

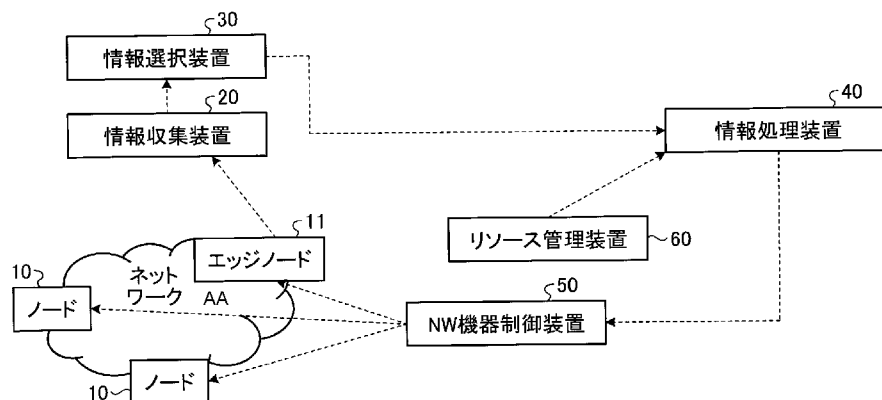
(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 藤原 貴之 (FUJIWARA, Takayuki); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 N T T 知的財産センタ内 Tokyo (JP). 西木 雅幸 (NISHIKI, Masayuki); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 N T T 知的財産センタ内 Tokyo (JP). 武田 知典 (TAKEDA, Tomonori); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 N T T 知的財産センタ内 Tokyo (JP).

(74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎の門三井ビルディング Tokyo (JP).

(54) Title: COMMUNICATION SYSTEM AND COMMUNICATION METHOD

(54) 発明の名称: 通信システム及び通信方法



- 10 Node
- 11 Edge node
- 20 Information collection device
- 30 Information selection device
- 40 Information processing device
- 50 NW apparatus control device
- 60 Resource management device
- AA Network

(57) Abstract: A communication system comprises: an information collection device (20) which collects, from an edge node (11), INT information attached to a packet, indicating passing time of an input port and passing time of an output port of each node (10); an information selection device (30) which selects, among the collected pieces of time stamp information, time stamp information corresponding to a selection condition; an information processing device (40) which performs prescribed calculation processing on the basis of the selected time stamp information; and an NW apparatus

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

control device (50) which controls a network apparatus on the basis of the result of the calculation by the information processing device (40).

(57) 要約: 通信システムは、エッジノード(11)から、パケットに付与された各ノード(10)の入力ポートの経路時間と出力ポートの経路時間とを示すINT情報を収集する情報収集装置(20)と、収集されたタイムスタンプ情報のうち、選択条件に応じたタイムスタンプ情報を選択する情報選択装置(30)と、選択されたタイムスタンプ情報を基に所定の計算処理を行う情報処理装置(40)と、情報処理装置(40)による計算結果を基にネットワーク機器を制御するNW機器制御装置(50)と、を有する。

明 細 書

発明の名称：通信システム及び通信方法

技術分野

[0001] 本発明は、通信システム及び通信方法に関する。

背景技術

[0002] ネットワーク管理のために、ネットワーク状態を監視する方法が提案されている。例えば、従来からよく用いられている方法として、S N M P (Simple Network Management Protocol) がある（非特許文献 1 参照）。

[0003] この S N M P は、管理及び監視を行う S N M P サーバが、監視下の装置に、該装置の状態情報の送信リクエストを送り、監視下の装置が自装置の状態を S N M P サーバにレスポンスする。この監視下の装置は、S N M P サーバによるリクエストがある度に S N M P サーバにレスポンスを行う。このため、S N M P では、S N M P サーバの監視下の装置の状態しか把握できない。具体的には、S N M P では、パケットロスが発生したことはでは判別できるものの、いずれのユーザのいずれのパケットに影響が出たか、までを特定することはできない。

[0004] また、他の方法として、Streaming Telemetry がある（非特許文献 2 参照）。Streaming Telemetry では、監視下の装置に、予め、所定時間毎に状態通知を送信するように設定を行う。このため、監視下の各装置は、レスポンス相当の内容を定期的に送信する。この結果、Streaming Telemetry では、S N M P と比して、各装置の負荷が軽減できる。

[0005] これに対し、I N T (In-band Network Telemetry) は、上記の方法が用いた状態通知専用のメッセージの送信に変えて、監視下の各装置が、ユーザが送信するパケットデータ自体に専用のヘッダを埋め込む処理を行う（例えば、非特許文献 3 参照）。

[0006] I N T では、パケット単位で情報を付与するため、上記の S N M P や Streaming Telemetry と異なり、パケットロスや遅延が発生した場合に、どのパケ

ットが影響を受けたかを特定することが可能である。そして、サーバ及び監視下の装置が、リクエスト及びレスポンスの情報を個別に作成するSNMPと比して、INTでは、サーバ及び監視下の装置の負担は軽減できるものと考えられる。このように、INTは、従来のSNMP、Streaming Telemetryと状態情報の通知方法が異なり、さらに、監視対象の粒度を小さくすることができる。

先行技術文献

非特許文献

[0007] 非特許文献1：A Simple Network Management Protocol, [online], [平成30年12月10日検索], インターネット<URL : <https://tools.ietf.org/html/rfc1157>>

非特許文献2：テレメトリデータのストリーミング, [online], [平成30年12月10日検索], インターネット<URL : https://www.cisco.com/c/ja_jp/td/docs/rt/wanaggregationinternetedge/nwconvergencesystem5000/cg/011/Telemetry-Config-Guide/Telemetry-Config-Guide_chapter_00.pdf>

非特許文献3：In-band Network Telemetry (INT), [online], [平成30年12月10日検索], インターネット<<https://p4.org/assets/INT-current-spec.pdf>>

発明の概要

発明が解決しようとする課題

[0008] しかしながら、INTを用いた場合、SNMP等を用いる場合よりも装置の状態及び通信品質の状況を詳細に監視できる反面、取得する情報量が膨大になる。この結果、取得する全INT情報を用いた場合、膨大な情報量を保持することが必要になるという問題や、情報処理に要する計算時間が長大になるという問題が生じ、ネットワークの状態監視を適切かつ効率的に実行することが難しい場合があった。

[0009] 本発明は、上記に鑑みてなされたものであって、ネットワークの状態監視を適切かつ効率的に実行することができる通信システム及び通信方法を提供することを目的とする。

課題を解決するための手段

[0010] 上述した課題を解決し、目的を達成するために、本発明に係る通信システムは、自装置を経由するパケットに対して自装置の識別情報とともに入力ポートの経由時間と出力ポートの経由時間とを示すタイムスタンプ情報を付与する転送装置から、タイムスタンプ情報を収集する収集部と、収集部によって収集されたタイムスタンプ情報のうち、選択条件に応じたタイムスタンプ情報を選択する選択部と、選択部によって選択されたタイムスタンプ情報を基に所定の計算処理を行う処理部と、処理部による計算結果を基にネットワーク機器を制御する制御部と、を有することを特徴とする。

発明の効果

[0011] 本発明によれば、ネットワークの状態監視を適切かつ効率的に実行することができる。

図面の簡単な説明

[0012] [図1]図1は、実施の形態1における通信システムの構成の一例を示す図である。

[図2]図2は、パケットに付与されたINT情報のデータ構成の一例を示す図である。

[図3]図3は、図1に示す通信システムにおける通信処理の流れについて説明する図である。

[図4]図4は、INT情報の選択方法を説明する図である。

[図5]図5は、INT情報の選択方法を説明する図である。

[図6]図6は、実施の形態1に係る通信処理の処理手順を示すシーケンス図である。

[図7]図7は、実施の形態2における通信システムの構成の一例を示す図である。

[図8]図8は、図7に示すデータベース（DB）が有する計算結果のデータ構成の一例を示す図である。

[図9]図9は、図7に示す通信システムにおける通信処理の流れについて説明する図である。

[図10]図10は、実施の形態2に係る通信処理の処理手順を示すシーケンス図である。

[図11]図11は、図10に示す比較処理の処理手順を示すフローチャートである。

[図12]図12は、プログラムが実行されることにより、情報収集装置、情報選択装置、情報処理装置、NW機器制御装置及び比較装置が実現されるコンピュータの一例を示す図である。

発明を実施するための形態

[0013] 以下、図面を参照して、本発明の一実施形態を詳細に説明する。なお、この実施の形態により本発明が限定されるものではない。また、図面の記載において、同一部分には同一の符号を付して示している。

[0014] [実施の形態1]

本発明の実施の形態1について説明する。図1は、実施の形態1における通信システムの構成の一例を示す図である。

[0015] 実施の形態1に係る通信システムは、図1に示すノード10及びエッジノード11を有するネットワークの状態を管理するシステムである。実施の形態1に係る通信システムは、情報収集装置20（収集部）、情報選択装置30（選択部）、情報処理装置40（情報処理部）及びNW機器制御装置50（制御部）を有する。

[0016] ノード10及びエッジノード11は、スイッチやルータ等の転送装置である。ノード及びエッジノード11は、光ファイバやLAN（Local Area Network）ケーブル等で接続され、ネットワークを構成する。エッジノード11は、ネットワークの出口に位置するノードである。

[0017] 各ノード10及びエッジノード11は、INT対応装置であり、自装置の

状態を通信システムに通知するために、INTを利用している。自装置の状態として、故障の有無、輻輳状態の有無等がある。ノード10及びエッジノード11は、自装置を経由するパケットに対して、INT情報を付与する。

INT情報は、自装置の識別情報とともに、自装置の入力ポートの経由時間と出力ポートの経由時間とを示すタイムスタンプ情報を含む。なお、実施の形態1では、INT情報を例に説明したが、各ノード10及びエッジノード11は、同様の仕組みで、タイムスタンプ情報とは異なるデータを、自装置を経由したパケットに付与することも可能である。

[0018] 図2は、パケットに付与されたINT情報のデータ構成の一例を示す図である。ここでは、ネットワーク内に存在するノード数は、全てINT対応装置であると仮定し、その数をMとする。図2に示すように、各ノード10及びエッジノード11から、INT情報として、例えば、各ノード10及びエッジノード11のingressポートとegressポートとにおけるタイムスタンプが付与される。

[0019] エッジノード11は、各パケットに付与されたINT情報をコピーし、情報収集装置20に転送する。ネットワークのエッジに存在する全てのノードは、いずれかのフローの終端点となり得るため、全てのエッジノード11は、この機能を有する。なお、エッジノード11は、INTヘッダの情報のみを情報収集装置20に転送してもよいし、ペイロードを含む情報を情報収集装置20に転送してもよい。

[0020] 情報収集装置20は、各ノード10及びエッジノード11から、INT情報を収集する。情報収集装置20は、エッジノード11から収集したINT情報を情報選択装置30へ転送する。図1の例では、情報収集装置20が、全てのエッジからの情報を集中的に管理しているものとする。なお、収集した情報の管理方法は、分散的または階層的であってもよい。

[0021] 情報選択装置30は、情報収集装置20によって収集されたINT情報のうち、選択条件に応じたINT情報を選択する。例えば、情報選択装置30は、特定のノード10及びエッジノード11を経由したパケットに付与され

た I N T 情報を選択する。情報選択装置 3 0 は、選択した I N T 情報を、情報処理装置 4 0 へ転送する。情報選択装置 3 0 は、選択条件の変更指示を受けた場合に、この変更指示に従って選択条件を変更する。

[0022] 情報処理装置 4 0 は、ネットワークの残帯域等のリソースを管理するリソース管理装置 6 0 と接続する。情報処理装置 4 0 は、情報選択装置 3 0 から得た I N T 情報と、必要に応じて取得したリソース管理装置 6 0 による入力情報とを基に、所定の計算処理を行う。例えば、情報処理装置 4 0 は、I N T 情報を基に、パケットの遅延時間を計算する。なお、情報処理装置 4 0 が接続する装置は、一例であり、リソース管理以外の他の機能を有する装置であってもよい。情報処理装置 4 0 は、計算結果を、NW機器制御装置 5 0 に出力する。

[0023] NW機器制御装置 5 0 は、情報処理装置 4 0 による計算結果を基に、ネットワーク機器（例えば、ノード）を制御する。

[0024] [通信処理の流れ]

次に、図 1 に示す通信システムにおける通信処理の流れについて説明する。図 3 は、図 1 に示す通信システムにおける通信処理の流れについて説明する図である。なお、図 3 における矢印 F 1 は、ある通信のフローを示す。ここで、フローとは、5-tuple (in/out のポート番号、in/out の I P アドレス、プロトコル番号) が等しいパケットの集合である。

[0025] 図 3 に示すように、情報収集装置 2 0 は、エッジノード 1 1 から、I N T 情報の収集を行う（図 3 の (1) 参照）。この際、情報収集装置 2 0 は、全パケットの I N T 情報を収集してもよく、また、一部のパケットの I N T 情報を収集してもよい。情報収集装置 2 0 は、収集した I N T 情報を情報選択装置 3 0 に転送する。情報選択装置 3 0 は、選択条件に応じた I N T 情報を選択し、選択した I N T 情報を情報処理装置 4 0 に転送する（図 3 の (2) 参照）。

[0026] 情報処理装置 4 0 は、情報選択装置 3 0 から転送された I N T 情報を基に、所定の計算処理を行う。この際、情報処理装置 4 0 は、リソース管理装置

60によるその他の有用な情報（例えば、リソース管理情報）を用いてもよい。

[0027] 例えば、情報処理装置40は、リソースが逼迫しているノード10及びエッジノード11を特定するための計算を行う。図2に示すように、INT情報は、ノード10及びエッジノード11のingressポートとegressポートにおけるタイムスタンプ情報を含む。情報処理装置40は、これらのタイムスタンプ情報に用いて、例えば、ネットワークに入ってから出てくるまでの間の遅延時間を計算することが可能である。具体的には、情報処理装置40は、（1）式を用いて、遅延時間の計算対象であるパケットkがネットワークに入ってから出てくるまでの遅延時間 T_d^k を計算する。

[0028] [数1]

$$T_d^k = T_{out}^{kM} - T_{in}^{k1} \quad \dots (1)$$

[0029] また、情報処理装置40は、これらのタイムスタンプ情報に用いて、ノード10及びエッジノード11ごとに、遅延時間の計算対象であるパケットkが遅延した時間を計算することができる。具体的には、情報処理装置40は、（2）式を用いて、遅延時間の計算対象であるパケットkが装置iで遅延した時間 T_{di}^k を計算し、特定の装置だけ遅延が大きくなっているかどうかを確認することができる。

[0030] [数2]

$$T_{di}^k = T_{out}^{ki} - T_{in}^{ki} \quad \dots (2)$$

[0031] 続いて、情報処理装置40は、計算結果を、NW機器制御装置50に出力する（図3の（3）参照）。NW機器制御装置50は、情報処理装置40による計算結果を基に、ネットワーク機器（ノード10及びエッジノード11）を制御する。

[0032] さらに、オペレータ等が、情報処理装置40の計算結果の精度が十分ではないと判断する場合がある。この場合には、オペレータ等による操作によって、情報選択装置30におけるINT情報の選択方法の変更が指示される（

図3の(4)参照)。情報選択装置30は、選択方法の変更指示を受けると、INT情報の選択方法を、指示された方法に変更する。ここで、計算結果の精度が悪い状況とは、例えば、情報処理装置40による計算よりも先に、何らかの外的要因によって輻輳発生箇所が判明するような状況である。

[0033] [INT情報の選択方法]

ネットワーク内の輻輳を検知する上で、エッジノード11が全てのINT情報を情報収集装置20に転送し、これら全てのINT情報を情報処理装置40における計算対象とした場合には、膨大な情報量を保持することが必要になり、情報処理に要する計算時間が実用的な時間で完了できなくなる。本実施の形態1では、情報選択装置30が、収集されたINT情報から、計算対象のINT情報を選択している。そこで、この情報選択装置30によるINT情報の選択方法について説明する。図4及び図5は、INT情報の選択方法を説明する図である。

[0034] 例えば、選択方法として、経由するフローが最も多い装置(ノード)から上位いくつかまでの装置で付与されたINT情報を計算対象として選択する方法(第1の選択方法)がある。図4の例の場合、経由するフロー数の多いノードは、ノード1, 2, 5, 6である。このため、情報選択装置30は、第1の選択方法が指示された場合、ノード1, 2, 5, 6においてパケットに付与されたINT情報を選択する。

[0035] また、選択方法として、より数の多い装置を経由したフローのパケットに付与されたINT情報を計算対象として選択する方法(第2の選択方法)がある。図4のフローF2~F4のうち、より数の多いノードを網羅するフローは、フローF3である。このため、情報選択装置30は、第2の選択方法が指示された場合、フローF3のパケットに付与されたINT情報を選択する。

[0036] 情報選択装置30に第1の選択方法が設定されており、オペレータ等によって、第2の選択方法への変更が指示された場合には、情報選択装置30は、第2の選択方法を用いて、以降のINT情報選択処理を行う。

[0037] ここで、第1及び第2の選択方法を用いて、計算対象のINT情報を制限して輻輳箇所を算出しようとする、全てのINT情報を用いる場合と比べて計算結果に誤差が生じることが考えられる。

[0038] 例えば、INT情報による遅延発生の発見や遅延箇所の特定よりも前に、SNMPによる定期監視やユーザからの申告等により、遅延の発生や遅延箇所が判明してしまうことが考えられる。このような場合、計算結果の精度を上げるために、情報選択装置30におけるINT情報の選択方法を変更する必要がある。

[0039] この場合、外的要因を把握したオペレータは、選択方法の変更指示を行い、これにともない、情報選択装置30は、指示された方法に、選択方法を変更する。そして、情報選択装置30は、変更した選択方法を用いてINT情報を選択し、この結果、情報処理装置40に入力するINT情報が変更される。

[0040] 例えば、INT情報を確認するノード10及びエッジノード11を限定する選択方法であった場合に、それ以外のノード10及びエッジノード11の負荷が高くなる等の理由で輻輳が発生させていることが推定より先に判明した場合を例に説明する。この場合、オペレータ等は、故障しているノード10及びエッジノード11を選択対象に含む選択方法に変更を指示することによって、情報選択装置30が、故障しているノード10及びエッジノード11を経由したパケットに付与されたINT情報を選択できるようにする。

[0041] 具体的には、フロー数の多いノードを上位4つだけ監視対象として選択する選択方法が適用されている場合を例に説明する。図5の例では、ノード1, 2, 5, 6が初期の監視対象となる。このため、情報選択装置30は、これらのノード1, 2, 5, 6で付与されたINT情報を選択し、情報処理装置40の入力情報として転送する。

[0042] ここで、ノード4で輻輳が発生していることが判明した場合（図5の（1）参照）について説明する。この場合には、オペレータ等の操作によって、輻輳が発生したノード4を含むように、監視対象のノードを、ノード4、1

、 2、 5 に変更する指示が情報選択装置 30 に通知される。このため、情報選択装置 30 は、これらのノード 4、 1、 2、 5 で付与された I N T 情報を選択し、情報処理装置 40 の入力情報として転送する。

[0043] [通信処理の処理手順]

次に、実施の形態 1 に係る通信処理の処理手順について説明する。図 6 は、実施の形態 1 に係る通信処理の処理手順を示すシーケンス図である。

[0044] 図 6 に示すように、ネットワークの出口に位置するエッジノード 11 が、各パケットに付与された I N T 情報をコピーし、情報収集装置 20 に転送する（ステップ S 1）。通信システムでは、情報収集装置 20 は、エッジノード 11 から転送された I N T 情報を収集し（ステップ S 2）、収集した I N T 情報を情報選択装置 30 へ転送する（ステップ S 3）。

[0045] 情報選択装置 30 は、情報収集装置 20 によって収集された I N T 情報のうち、選択条件に応じた I N T 情報を選択する（ステップ S 4）。情報選択装置 30 は、選択した I N T 情報を、情報処理装置 40 へ転送する（ステップ S 5）。

[0046] 情報処理装置 40 は、必要に応じてリソース管理装置 60 からリソース情報を取得する（ステップ S 6）。そして、情報処理装置 40 は、情報選択装置 30 による選択後の I N T 情報と、リソース情報とを基に、所定の計算処理を行う（ステップ S 7）。情報処理装置 40 は、計算結果を、例えば、N W機器制御装置 50 に出力する（ステップ S 8）。N W機器制御装置 50 は、情報処理装置 40 による計算結果を基に、ノード 10 及びエッジノード 11 を制御する（ステップ S 9～S 11）。

[0047] そして、入力装置に、オペレータ等による選択方法の変更指示が入力され（ステップ S 12）、選択条件の変更指示が通知された場合には（ステップ S 13）、情報選択装置 30 は、この変更指示に従って選択条件を変更する（ステップ S 14）。

[0048] [実施の形態 1 の効果]

このように、本実施の形態 1 に係る通信システムでは、エッジノード 11

から I N T 情報を収集し、収集した I N T 情報のうち、選択条件に応じたタイムスタンプ情報を選択し、選択した I N T 情報を基に所定の計算処理を行って、N W 機器を制御する。すなわち、本実施の形態 1 に係る通信システムでは、エッジノード 1 1 から収集した全ての I N T 情報を計算処理に用いるのではなく、収集した I N T 情報から所定の選択方法にしたがって選択された I N T 情報のみを用いて計算処理を行う。

[0049] したがって、本実施の形態 1 に係る通信システムによれば、収集した I N T 情報の全てを、計算処理のために保持する必要もない。また、本実施の形態 1 に係る通信システムによれば、選択された一部の I N T 情報を用いて計算処理を行うため、全 I N T 情報を用いた場合と比して、計算時間を短縮することができ、効率化を図ることができる。

[0050] また、実施の形態 1 にかかる通信システムでは、例えば、情報選択装置 3 0 の計算結果の精度が悪い場合など、計算結果の精度を上げるために、外部からの操作によって、I N T 情報の選択方法の変更が可能である。したがって、実施の形態 1 に係る通信システムは、ネットワークの状況に応じて、I N T 情報の選択方法を柔軟に変更でき、情報処理装置 4 0 の計算処理を適切に実行することができる。

[0051] このため、本実施の形態 1 によれば、計算処理に用いる I N T 情報をネットワーク状況に応じて選択することによって、ネットワークの状態監視を適切かつ効率的に実行することができる。

[0052] [実施の形態 2]

次に、実施の形態 2 について説明する。図 7 は、実施の形態 2 における通信システムの構成の一例を示す図である。

[0053] 図 7 に示すように、実施の形態 2 に係る通信システムでは、情報処理装置 4 0 が、過去の計算結果を記憶する D B 4 1 を有する。図 8 は、図 7 に示す D B 4 1 が有する計算結果のデータ構成の一例を示す図である。図 8 に示すように、情報処理装置 4 0 は、過去の計算結果として、計算開始時間及び各入力情報とに対応付けて各計算結果を保持する。例えば、通番「2」は、「2

018/10/03」に、入力1「D」、入力2「E」の情報を用いて計算を開始した計算結果「F」が記載されている。

[0054] そして、図7に示すように、実施の形態2に係る通信システムは、実施の形態1に係る通信システムと比して、さらに、比較装置280（比較部）を有する。比較装置280は、情報収集装置20、情報選択装置30、情報処理装置40及びリソース管理装置60と接続する。

[0055] 比較装置280は、情報収集装置20によって収集された全てのINT情報、及び、リソース管理装置60によるリソース情報等を用いて、所定の計算処理を行う。比較装置280は、情報処理装置40と同じ計算ロジックを内包する。

[0056] 比較装置280は、該計算結果及び情報処理装置40による計算結果の誤差を求める。そして、比較装置280は、求めた誤差と、予め設定された閾値との比較結果を基に、選択条件の変更を情報選択装置30に指示する。例えば、比較装置280は、求めた誤差が閾値を超えた場合に、情報選択装置30に選択方法の変更を指示する。この閾値は、例えば、オペレータ等の操作によって予め設定される。また、この閾値は、オペレータ等の操作によって、変更も可能である。

[0057] 例えば、情報選択装置30に第1の選択方法が適用されていた場合に、求めた誤差が閾値を超えた場合について説明する。この場合、比較装置280は、情報選択装置30に、第2の選択方法への変更を指示し、より数の多い装置を経由したフローのINT情報を取得して、多くの装置に対して状況を求めることができるようにする。

[0058] [通信処理の流れ]

次に、図7に示す通信システムにおける通信処理の流れについて説明する。図9は、図7に示す通信システムにおける通信処理の流れについて説明する図である。

[0059] 図9に示すように、情報収集装置20は、エッジノード11から、INT情報の収集を行い（図9の（1）参照）、情報選択装置30は、選択条件に

応じた I N T 情報を選択し、選択した I N T 情報を情報処理装置 4 0 に転送する（図 9 の（2-1）参照）。また、情報収集装置 2 0 は、収集した全ての I N T 情報を、比較装置 2 8 0 に転送する（図 9 の（2-2）参照）。

[0060] 情報処理装置 4 0 は、情報選択装置 3 0 から転送された I N T 情報を基に、所定の計算処理を行い、N W 機器制御装置 5 0 に加え、矢印 Y 1 に示すように、比較装置 2 8 0 にも出力する（図 9 の（3）参照）。N W 機器制御装置 5 0 は、情報処理装置 4 0 による計算結果を基に、ネットワーク機器を制御する。

[0061] そして、比較装置 2 8 0 は、情報収集装置 2 0 が収集した全 I N T 情報を用いて、情報処理装置 4 0 と同じ計算処理を実施する（図 9 の（4）参照）。この際、比較装置 2 8 0 は、比較装置 2 8 0 は、計算時間を限らずに、例えば遅延箇所を算出する。

[0062] 続いて、比較装置 2 8 0 は、情報処理装置 4 0 の D B 4 1 から、自装置の計算開始と同じ計算開始時間の入力と計算結果の組合せとを取得し、計算結果の誤差を確認する。例えば、比較装置 2 8 0 は、計算開始時間「2018/10/03」の計算処理を実施した場合、D B 4 1 が記憶するテーブル（図 8 参照）、計算開始時間「2018/10/03」の入力（入力 1 「D」、入力 2 「E」）及び結果「F」の情報を取得する（図 8 の行 R 2 参照）。そして、比較装置 2 8 0 は、自装置の計算結果と、取得した計算結果「F」とを比較し、誤差を求める。

[0063] そして、比較装置 2 8 0 は、求めた誤差が閾値を超過した場合には、情報選択装置 3 0 に、選択方法の変更を指示する（図 9 の（5）参照）。この指示に従い、情報選択装置 3 0 は、選択方法を変更する。

[0064] [通信処理の処理手順]

次に、実施の形態 1 に係る通信処理の処理手順について説明する。図 1 0 は、実施の形態 2 に係る通信処理の処理手順を示すシーケンス図である。

[0065] 図 1 0 に示すステップ S 2 1 ~ S 2 3 は、図 6 に示すステップ S 1 ~ S 3 と同じ処理である。実施の形態 2 では、情報収集装置 2 0 は、収集した全ての I N T 情報を、比較装置 2 8 0 へ転送する（ステップ S 2 4）。図 1 0 に

示すステップS 2 5 ～ S 3 2 は、図 6 に示すステップ S 4 ～ S 1 1 と同じ処理である。

[0066] 比較装置 2 8 0 は、情報処理装置 4 0 による計算結果を取得し（ステップ S 3 3）、全ての I N T 情報を用いた場合の計算結果と、取得した情報処理装置 4 0 による計算結果を比較して情報選択装置 3 0 における選択方法の変更を指示する比較処理を行う（ステップ S 3 4）。そして、比較装置 2 8 0 は、比較処理の結果に応じて、情報選択装置 3 0 に、I N T 情報の選択方法の変更を指示し（ステップ S 3 5）、情報選択装置 3 0 は、この変更指示に従って選択方法を変更する（ステップ S 3 6）。

[0067] [比較処理の処理手順]

次に、比較処理（ステップ S 3 4）の処理手順について説明する。図 1 1 は、図 1 0 に示す比較処理の処理手順を示すフローチャートである。

[0068] 図 1 1 に示すように、比較装置 2 8 0 は、情報収集装置 2 0 が収集した全 I N T 除法を取得する（ステップ S 4 1）。比較装置 2 8 0 は、全 I N T 情報を用いて所定の計算処理を行う（ステップ S 4 2）。そして、情報処理装置 4 0 による計算結果を取得して（ステップ S 4 3）、自装置で計算した結果と、情報処理装置 4 0 から取得した計算結果との誤差を求める（ステップ S 4 4）。

[0069] そして、比較装置 2 8 0 は、求めた誤差が閾値を超えたか否かを判定する（ステップ S 4 5）。比較装置 2 8 0 は、求めた誤差が閾値を超えたと判定した場合（ステップ S 4 5：Y e s）、情報選択装置 3 0 に選択方法の変更指示を通知し（ステップ S 4 6）、比較処理を終了する。一方、比較装置 2 8 0 は、求めた誤差が閾値を超えていないと判定した場合（ステップ S 4 5：N o）、比較処理を終了する。

[0070] [実施の形態 2 の効果]

このように、実施の形態 2 に係る通信システムは、比較装置 2 8 0 をさらに設け、選択方法の変更を、通信システム内部で自動的に実行する。このため、実施の形態 1 と比して、オペレータ等による指示の前に、I N T 情報の

選択方法を変更することができ、情報処理の計算処理の精度の維持を図ることができる。

[0071] なお、比較装置 280 は、自装置による計算結果及び情報処理装置 40 による計算結果の誤差が、閾値を超えた度合いに応じて、情報選択装置 30 に指示する選択方法を選択してもよい。例えば、比較装置 280 は、求めた誤差が閾値を超えており、求めた誤差と閾値との差が所定値以上である場合には、情報選択装置 30 に、第 2 の選択方法への変更を指示する。一方、比較装置 280 は、求めた誤差が閾値を超えており、求めた誤差と閾値との差が所定値未満である場合には、情報選択装置 30 に、第 2 の選択方法とは異なる第 3 の選択方法への変更を指示する。

[0072] [システム構成等]

図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各装置の分散・統合の具体的形態は図示のものに限られず、その全部又は一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的又は物理的に分散・統合して構成することができる。さらに、各装置にて行なわれる各処理機能は、その全部又は任意の一部が、CPU (Central Processing Unit) 及び当該 CPU にて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

[0073] また、本実施形態において説明した各処理のうち、自動的に行われるものとして説明した処理の全部又は一部を手動的に行なうこともでき、あるいは、手動的に行なわれるものとして説明した処理の全部又は一部を公知の方法で自動的行なうこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。

[0074] [プログラム]

図 12 は、プログラムが実行されることにより、情報収集装置 20、情報選択装置 30、情報処理装置 40、NW 機器制御装置 50 及び比較装置 28

0が実現されるコンピュータの一例を示す図である。コンピュータ1000は、例えば、メモリ1010、CPU1020を有する。また、コンピュータ1000は、ハードディスクドライバインタフェース1030、ディスクドライバインタフェース1040、シリアルポートインタフェース1050、ビデオアダプタ1060、ネットワークインタフェース1070を有する。これらの各部は、バス1080によって接続される。

[0075] メモリ1010は、ROM (Read Only Memory) 1011及びRAM (Random Access Memory) 1012を含む。ROM1011は、例えば、BIOS (Basic Input Output System) 等のブートプログラムを記憶する。ハードディスクドライバインタフェース1030は、ハードディスクドライブ1090に接続される。ディスクドライバインタフェース1040は、ディスクドライブ1100に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ1100に挿入される。シリアルポートインタフェース1050は、例えばマウス1110、キーボード1120に接続される。ビデオアダプタ1060は、例えばディスプレイ1130に接続される。

[0076] ハードディスクドライブ1090は、例えば、OS (Operating System) 1091、アプリケーションプログラム1092、プログラムモジュール1093、プログラムデータ1094を記憶する。すなわち、情報収集装置20、情報選択装置30、情報処理装置40、NW機器制御装置50及び比較装置280の各処理を規定するプログラムは、コンピュータにより実行可能なコードが記述されたプログラムモジュール1093として実装される。プログラムモジュール1093は、例えばハードディスクドライブ1090に記憶される。例えば、情報収集装置20、情報選択装置30、情報処理装置40、NW機器制御装置50及び比較装置280における機能構成と同様の処理を実行するためのプログラムモジュール1093が、ハードディスクドライブ1090に記憶される。なお、ハードディスクドライブ1090は、SSD (Solid State Drive) により代替されてもよい。

[0077] また、上述した実施形態の処理で用いられる設定データは、プログラムデータ1094として、例えばメモリ1010やハードディスクドライブ1090に記憶される。そして、CPU1020が、メモリ1010やハードディスクドライブ1090に記憶されたプログラムモジュール1093やプログラムデータ1094を必要に応じてRAM1012に読み出して実行する。

[0078] なお、プログラムモジュール1093やプログラムデータ1094は、ハードディスクドライブ1090に記憶される場合に限らず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ1100等を介してCPU1020によって読み出されてもよい。あるいは、プログラムモジュール1093及びプログラムデータ1094は、ネットワーク（LAN、WAN（Wide Area Network）等）を介して接続された他のコンピュータに記憶されてもよい。そして、プログラムモジュール1093及びプログラムデータ1094は、他のコンピュータから、ネットワークインタフェース1070を介してCPU1020によって読み出されてもよい。

[0079] 以上、本発明者によってなされた発明を適用した実施形態について説明したが、本実施形態による本発明の開示の一部をなす記述及び図面により本発明は限定されることはない。すなわち、本実施形態に基づいて当業者等によりなされる他の実施形態、実施例及び運用技術等は全て本発明の範疇に含まれる。

符号の説明

- [0080]
- 10 ノード
 - 11 エッジノード
 - 20 情報収集装置
 - 30 情報選択装置
 - 40 情報処理装置
 - 50 NW機器制御装置
 - 60 リソース管理装置

請求の範囲

- [請求項1] 自装置を経由するパケットに対して自装置の識別情報とともに入力ポートの経由時間と出力ポートの経由時間とを示すタイムスタンプ情報を付与する転送装置から、前記タイムスタンプ情報を収集する収集部と、
- 前記収集部によって収集されたタイムスタンプ情報のうち、選択条件に応じたタイムスタンプ情報を選択する選択部と、
- 前記選択部によって選択されたタイムスタンプ情報を基に所定の計算処理を行う処理部と、
- 前記処理部による計算結果を基にネットワーク機器を制御する制御部と、
- を有することを特徴とする通信システム。
- [請求項2] 前記選択部は、前記選択条件の変更指示を受けた場合に、前記変更指示に従って前記選択条件を変更することを特徴とする請求項1に記載の通信システム。
- [請求項3] 前記収集部によって収集されたタイムスタンプ情報を用いて前記所定の計算処理を行い、該計算結果及び前記処理部による計算結果の誤差と、予め設定された閾値との比較結果を基に、前記選択条件の変更を前記選択部に指示する比較部をさらに有することを特徴とする請求項1に記載の通信システム。
- [請求項4] 前記転送装置は、I N T (In-band Network Telemetry) 対応装置であり、
- 前記タイムスタンプ情報は、前記転送装置が付与したI N T 情報に含まれることを特徴とする請求項1～3のいずれか一つに記載の通信システム。
- [請求項5] 前記処理部は、前記I N T 情報を基に、前記パケットの遅延時間を計算することを特徴とする請求項4に記載の通信システム。
- [請求項6] 通信システムが実行する通信方法であって、

自装置を経由するパケットに対して自装置の識別情報とともに入力ポートの経由時間と出力ポートの経由時間とを示すタイムスタンプ情報を付与する転送装置から、前記タイムスタンプ情報を収集する収集工程と、

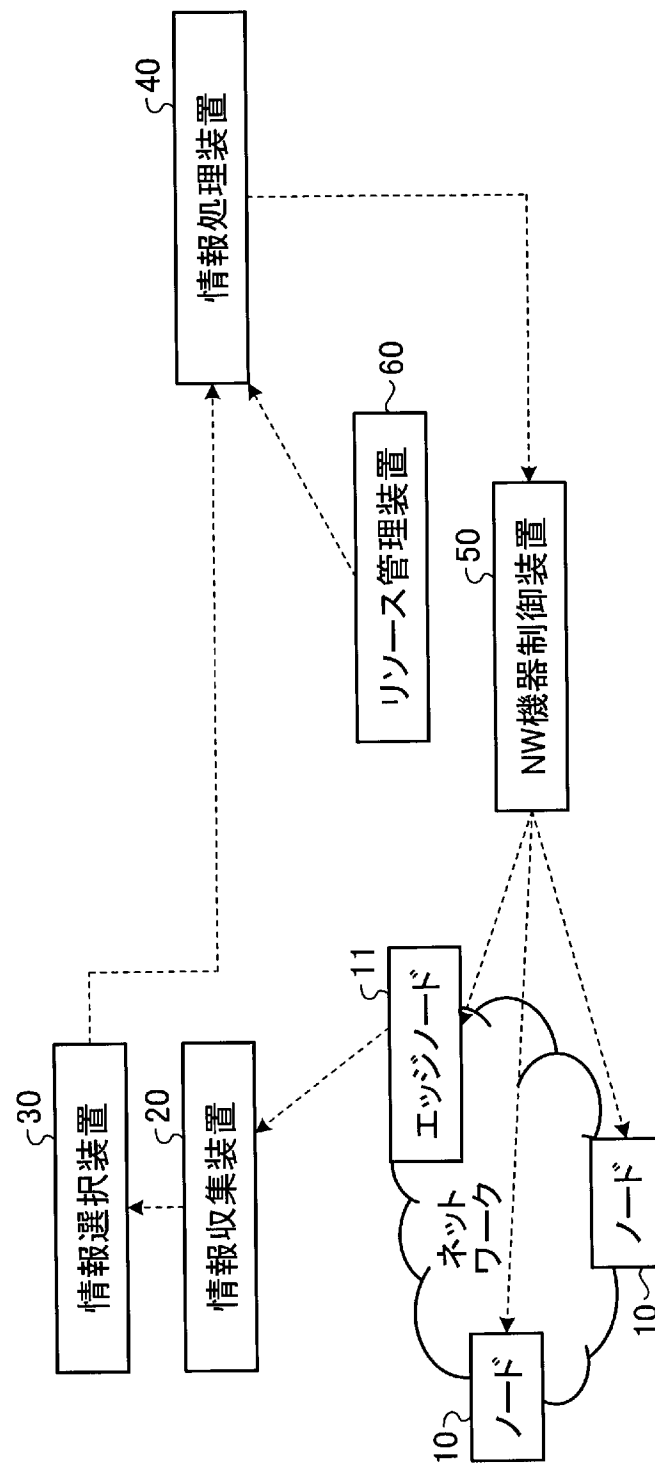
前記収集工程において収集されたタイムスタンプ情報のうち、選択条件に応じたタイムスタンプ情報を選択する選択工程と、

前記選択工程において選択されたタイムスタンプ情報を基に所定の計算処理を行う処理工程と、

前記処理工程における計算結果を基にネットワーク機器を制御する制御工程と、

を含んだことを特徴とする通信方法。

[図1]



[図2]

	D_{in}^1	D_{out}^1	D_{in}^2	D_{out}^2	...	D_{in}^M	D_{out}^M
パケットk	T_{in}^{k1}	T_{out}^{k1}	T_{in}^{k2}	T_{out}^{k2}	...	T_{in}^{kM}	T_{out}^{kM}

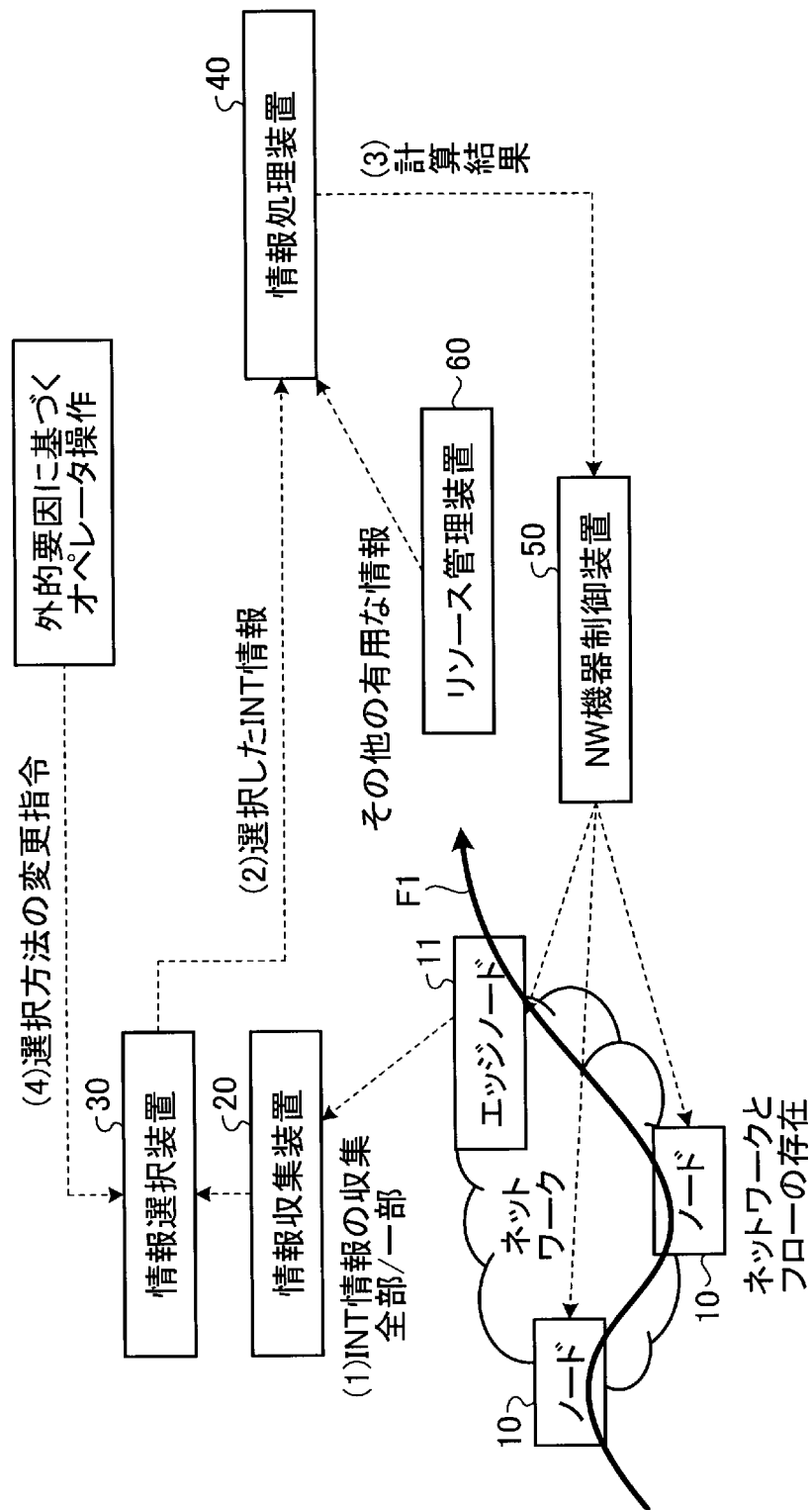
D_{in}^M : M番目のノードのingress

D_{out}^M : M番目のノードのegress

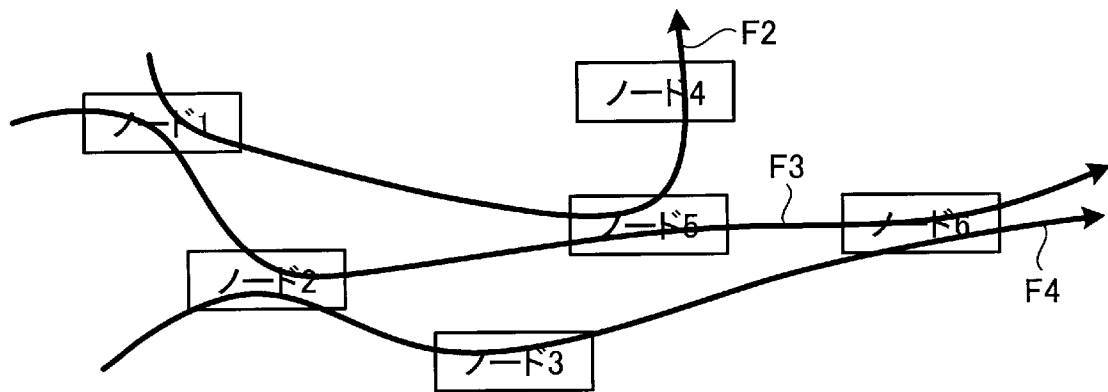
T_{in}^{kM} : k番目のパケットがM番目のノードのingressを通過した時刻情報

T_{out}^{kM} : k番目のパケットがM番目のノードのegressを通過した時刻情報

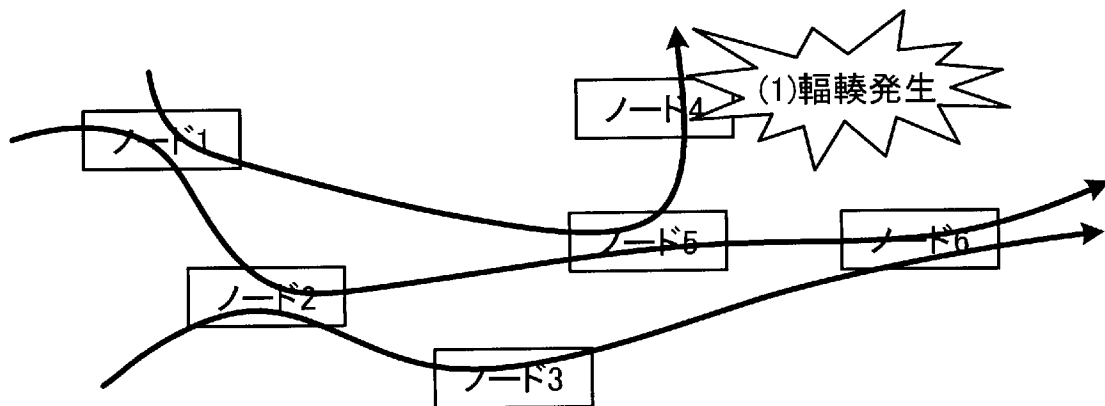
[図3]



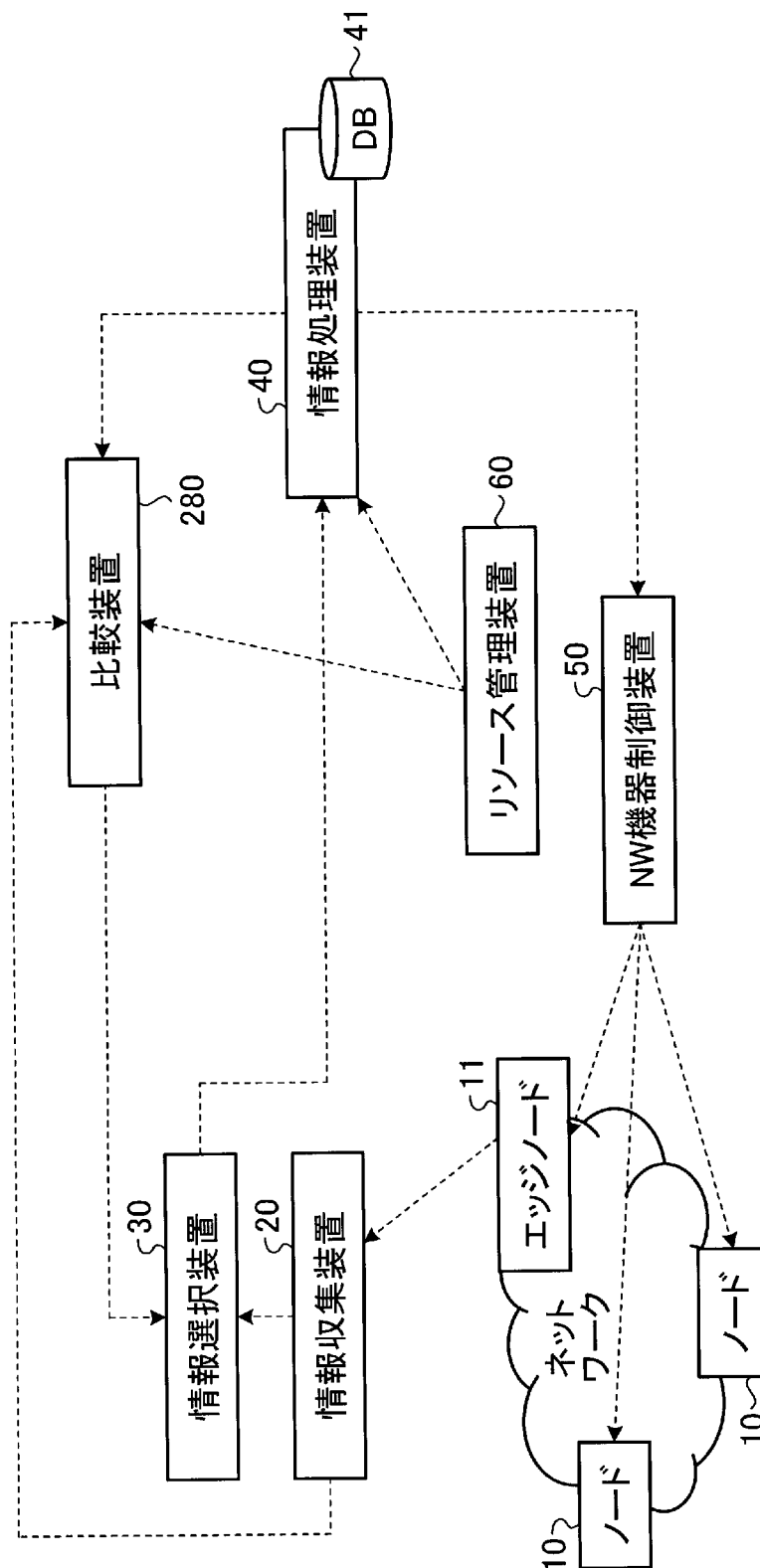
[図4]



[図5]



[図7]

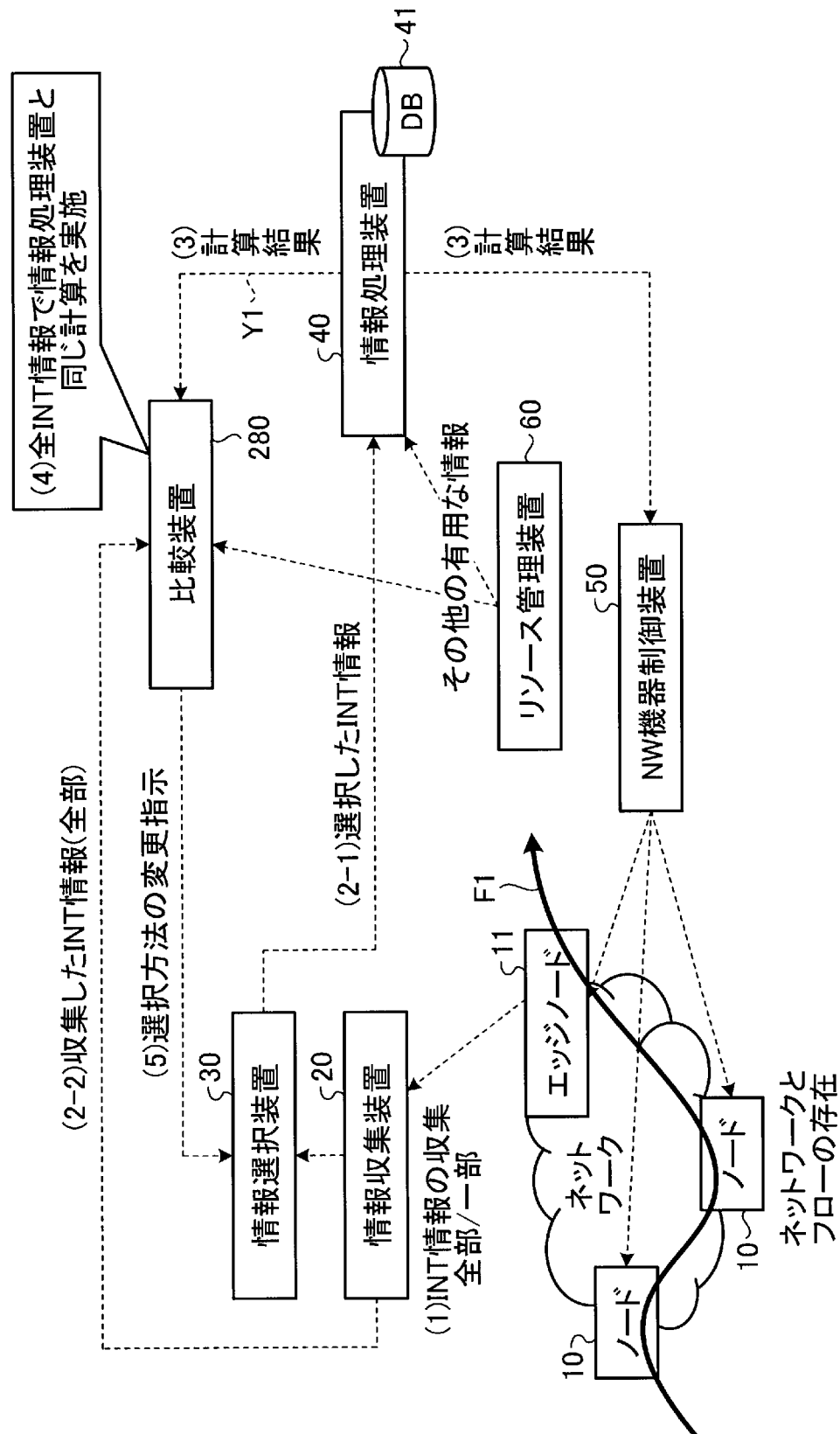


[図8]

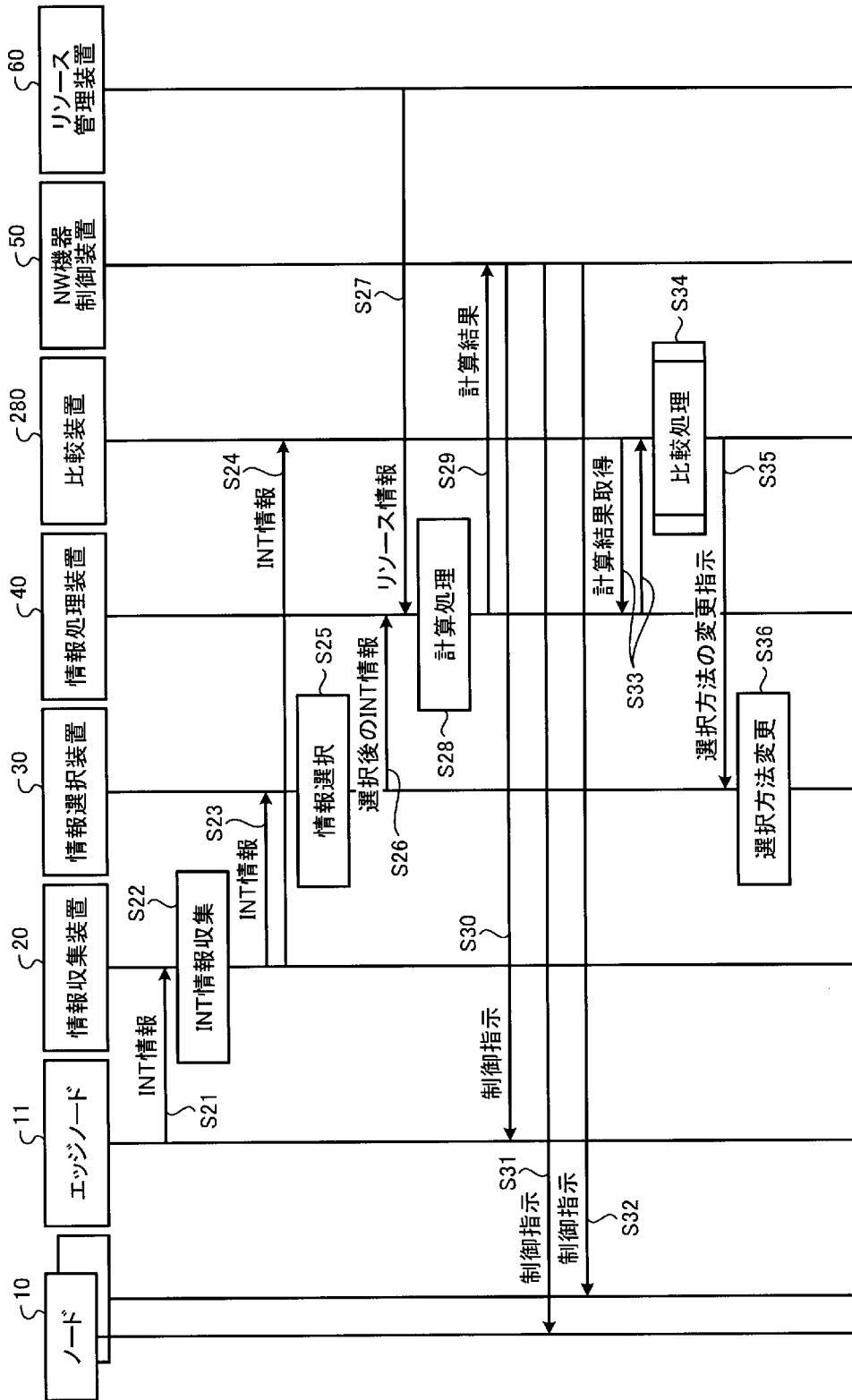
項番	計算開始時間	入力1	入力2	...	結果
1	2018/11/08 11:44	A	B	...	C
2	2018/10/03 04:30	D	E	...	F
⋮	⋮	⋮	⋮	⋮	⋮

} R2

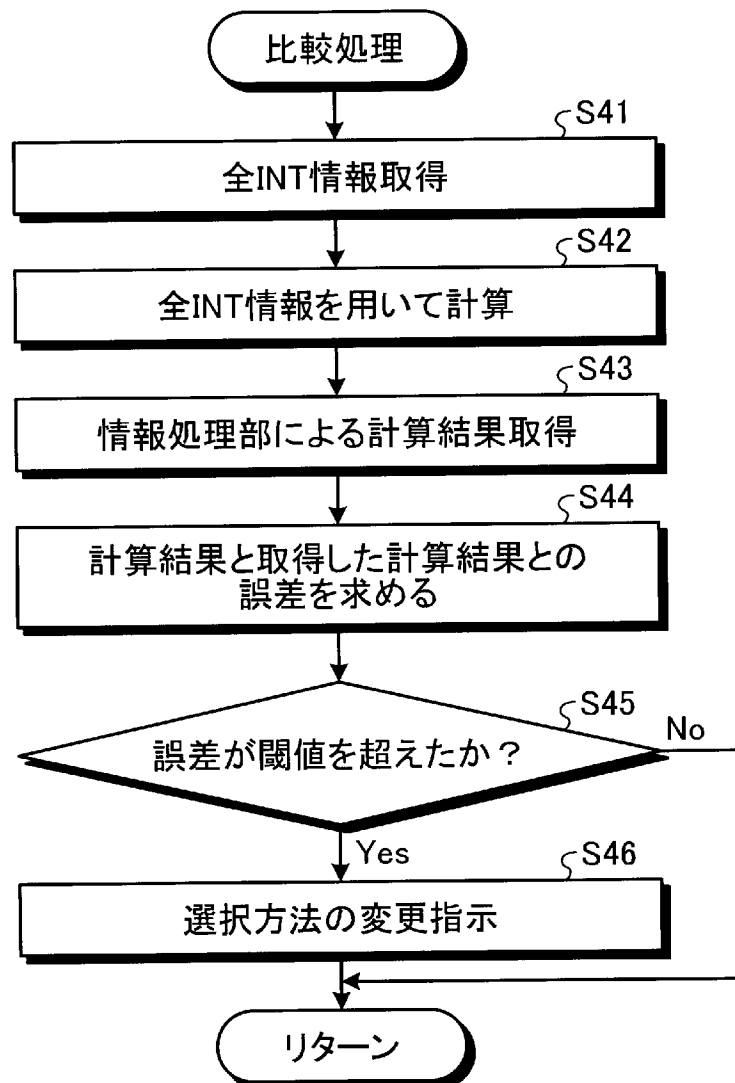
[図9]



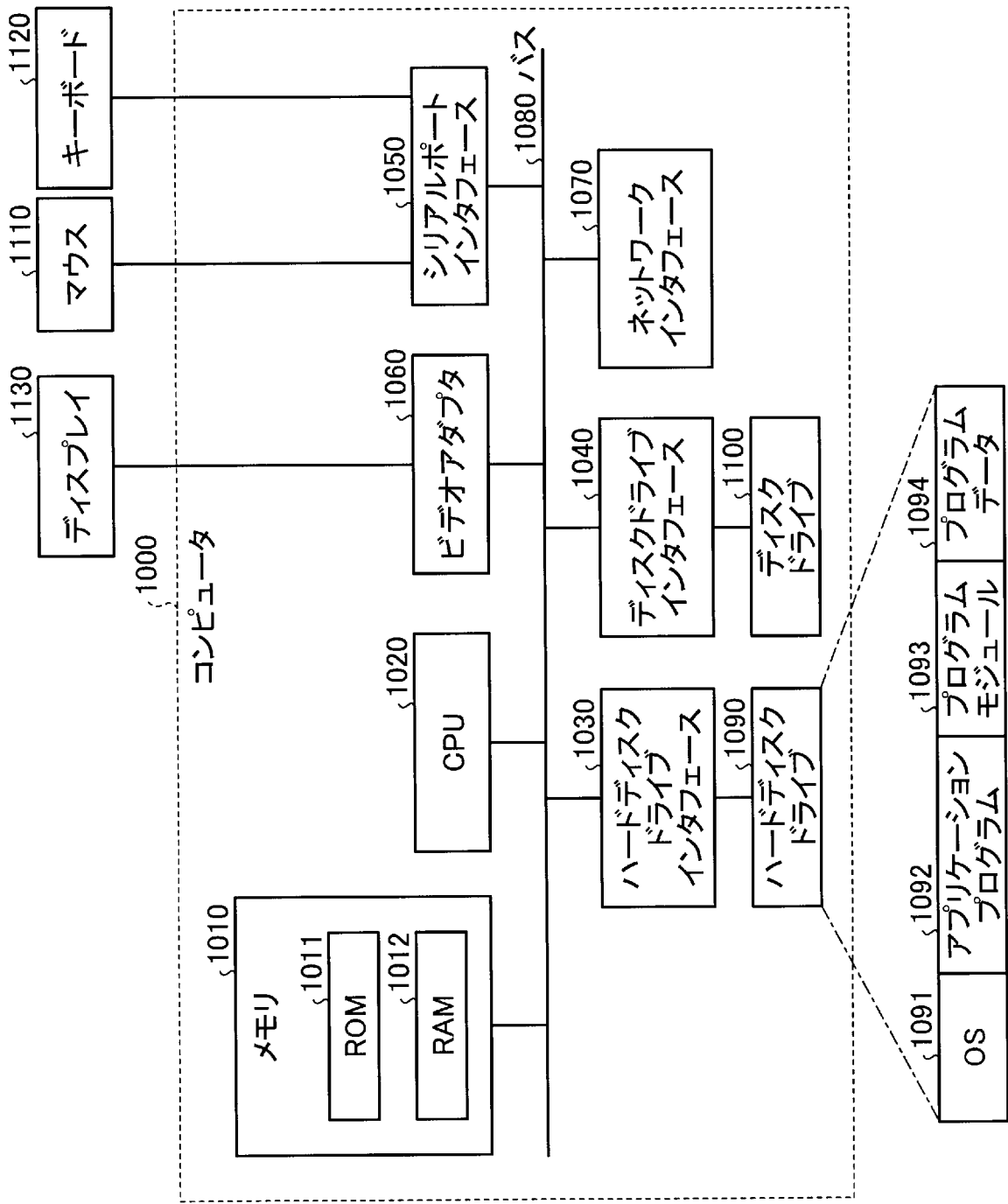
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/046289

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04L12/70 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H04L12/70

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	TU, N. V. et al., INTCollector: A high-performance collector for in-band network telemetry, 2018 14th International Conference on Network and Service Management (CNSM), IEEE, 24 December 2018, pp. 10-18	1-2, 4-6 3
A	MARVELL, Network telemetry solutions for data center and enterprise networks, White Paper [online], March 2018 [retrieved on 18 December 2019], Internet <URL:https://www.marvell.com/documents/8ni0aekscuc444ay31nn/>, pp. 1-14	1-6



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

19.12.2019

Date of mailing of the international search report

07.01.2020

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/046289

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KIM, C. H. et al., In-band network telemetry (INT) [online], June 2016 [retrieved on 18 December 2019], Internet <URL:https://p4.org/assets/INT-current-spec.pdf>, pp. 1-28	1-6
A	WO 2015/040624 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY, L. P.) 26 March 2015, paragraphs [017], [019], [026]-[033], fig. 1, 2 & US 2016/0226742 A1	1-6
A	US 9450846 B1 (CISCO TECHNOLOGY, INC.) 20 September 2016, column 9, lines 19-62, fig. 3 (Family: none)	1-6
P, X	藤原 貴之, In-band network telemetry を用いたネットワーク監視方式の一考察, 電子情報通信学会技術研究報告, vol. 119, no. 5, [online], 11 April 2019, pp. 85-90, NS2019-15, (FUJIWARA, Takayuki, A study on network monitoring method using in-band network telemetry, IEICE Technical Report)	1-2, 4-6

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. H04L12/70(2013.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. H04L12/70

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X A	Nguyen Van Tu et al., INTCollector: A High-performance Collector for In-band Network Telemetry, 2018 14th International Conference on Network and Service Management (CNSM), IEEE, 2018.12.24, P.10-18	1-2, 4-6 3
A	MARVELL, Network Telemetry Solutions for Data Center and Enterprise Networks, WHITE PAPER [オンライン], 2018.03 [検索日 2019.12.18], インターネット<URL: https://www.marvell.com/documents/8ni0aekscuc444ay3l1nn/ >	1-6

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」 同一パテントファミリー文献

国際調査を完了した日

19.12.2019

国際調査報告の発送日

07.01.2020

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

羽岡 さやか

5 X

3149

電話番号 03-3581-1101 内線 3596

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
	P. 1-14	
A	Changhoon Kim et al., In-band Network Telemetry (INT) [オンライン], 2016.06 [検索日 2019.12.18], インターネット<URL: https://p4.org/assets/INT-current-spec.pdf > P. 1-28	1-6
A	WO 2015/040624 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.) 2015.03.26, [017], [019], [026]-[033], FIG. 1, FIG. 2 & US 2016/0226742 A1	1-6
A	US 9450846 B1 (CISCO TECHNOLOGY, INC.) 2016.09.20, 第9欄 19-62行, FIG. 3 (ファミリーなし)	1-6
P, X	藤原 貴之 Takayuki FUJIWARA, In-band Network Telemetryを用いたネットワーク監視方式の一考察 A Study on Network Monitoring Method using In-band Network Telemetry, 電子情報通信学会技術研究報告 Vol. 119 No. 5 [online] IEICE Technical Report, 2019.04.11, 第119巻, P.85-90, NS2019-15	1-2, 4-6