

ドメインリーダーに基づく複製管理方式の評価と改良

中村 健二[†] 宮西 洋太郎^{††}
佐藤 文明^{†††} 水野 忠 則^{†††}

近年、ワイヤレス通信技術の進歩、ハードウェアの小型軽量化によって、携帯型のPCや個人情報管理機器がネットワークに接続されるようになってきた。このようなモバイル環境では、いつでもどこでもデータのやり取りを可能とするアプリケーションが開発されている。通信速度が有線に比べて遅いモバイル環境では、データ処理の応答時間を短縮するために、複製を用いることが望ましいと思われる。しかし、従来の複製管理アルゴリズムはいくつかの問題があった。この問題点を改善した複製管理アルゴリズムとして、ドメインリーダーアルゴリズムが提案されている。しかし、簡単な解析的な評価が行われているだけで、モバイル環境での利用についての十分な評価が行われていなかった。本論文では、ドメインリーダーの基本的なアルゴリズムの他に、いくつかのシステム状況を考慮して2つの拡張を提案し、シミュレーションによって評価した。さらに、ドメインリーダーの性能について、ある条件においては最適なりーダ数が存在することを示した。また、通信コストの比較によって、ドメインリーダー方式が従来のアルゴリズムよりモバイル環境に適していることを明らかにした。

Evaluation and Improvement of Replica Control Protocol Based on the Domain Leader Concept

KENJI NAKAMURA,[†] YOHTARO MIYANISHI,^{††} FUMIAKI SATO^{†††}
and TADANORI MIZUNO^{†††}

Recently, as wireless communication technologies and micro-electronics technologies make a great advance, portable PCs and personal information management tools have been networked. In such mobile environment, many applications offering data access to users in every place and every time have been developed. Since the communication speed and reliability in the mobile environment is lower than wired environment, data replication is useful technology. However, existing researches are not adequate to the mobile environment. The Domain Leader algorithm has been developed to overcome these problems. However the algorithm has not been evaluated in detail. The purpose of this paper is to evaluate the Domain Leader algorithm with many parameters, and to present some extended algorithms. We clarify that the Domain Leader algorithm has an optimal number of leaders in the specific configuration. We also present the suitability of the algorithm to the heterogeneous communication environment such as the mobile environment.

1. はじめに

近年、ネットワークや計算機技術の発達により通信速度が高速化し、CPUやディスクなどの資源の価格が大幅に低下した。これによって、ネットワークと低価格の資源を有効に使う分散データベースシステムが出現した。また、分散データベースシステム上のデー

タの移動性を提供するモバイルシステム^{1)~3)}の研究も多々行われるようになってきた。

その分散データベースシステムで用いるデータの信頼性と可用性を向上させるために、データの複製管理方式が研究されている。複製化されたデータの一貫性と正当性は、複製管理アルゴリズムによって保証される。我々は、すでに、大規模なシステムを前提とする複製管理方式としてドメインリーダー方式を提案している⁴⁾。従来までの研究では、ドメインリーダー方式の解析的な評価を行ったが、本論文では、シミュレーションとアクセスサイト数の比較、および通信コストによる詳細な評価を行う。また、特定の環境に適応した改良アルゴリズムについて述べる。さらに、モバイル

[†] 日本ヒューレット・パッカード株式会社
Hewlett-Packard Japan, Ltd.

^{††} 三菱電機株式会社
Mitsubishi Electric Corporation

^{†††} 静岡大学情報学部
Faculty of Information, Shizuoka University

環境などの均質でないネットワークにおいて本方式が有効であることを示す。

以下、本論文は2章において従来の複製管理方式の特徴と問題点について述べる。3章では、本論文で対象とするドメインリーダー方式について述べる。4章では、既存のドメインリーダー方式と、本論文で提案する特定の環境に適合する拡張アルゴリズムについてシミュレーションにより評価する。5章では、ドメインリーダー方式における、故障対策について述べる。6章は、ドメインリーダー方式をモバイル環境に適用した場合について、アクセスサイト数に基づく通信コストの比較によって評価した結果を示す。7章は本論文のまとめである。

2. 複製管理アルゴリズム

複製管理アルゴリズムは、複製データの一貫性と正当性を保証し、複数のサイトからなるトランザクションの同時実行制御を行う。複製管理技術は、分散オペレーティングシステムや分散データベースの研究で広く行われている。以下に、代表的な複製管理方式の特徴を述べる。

プライマリコピー：読み込みを実行する場合、1個のサイトしかアクセスしないので通信コストは最小となるが、書き込みを実行する場合、全サイトをアクセスするので、複製数が多いと、通信コストが高くなる。

定数合意：データのアクセス形態に応じて自由にアクセスするサイト数を変更することができるが、複製数が多いと、やはり複製数に比例してアクセス回数が増える⁵⁾。

階層型定数合意：サイトのアクセス回数は複製数を N として $N^{0.63}$ のアクセス回数となるため、効率的である。また、読み込みと書き込みのサイト数の比は、定数合意と同様にアプリケーションに対応して変更ができる⁶⁾。

三角格子：この方式は読み込みと書き込みのサイトアクセス数が $N^{0.5}$ と小さい。しかし、読み込みと書き込みのアクセス数の比は変更できない⁷⁾。

また、階層型定数合意と三角格子は、全サイトを論理的にツリー構造あるいは三角の格子のように構成する必要があるため、管理が複雑となるという問題がある。また、これらの従来よりの方法は、ネットワークやデータベースが均質であるという仮定で評価されており、通信速度が不均質なモバイル環境では有効とは限らない。

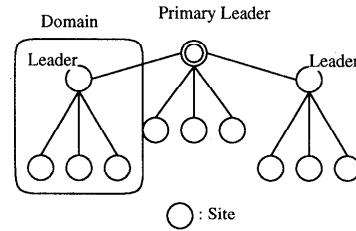


図1 ドメインリーダー方式のシステムモデル

Fig.1 System model of Domain Leader method.

3. ドメインリーダーアルゴリズム

本章では、文献^{4),8)~10)}にあるドメインリーダーアルゴリズムのプロトコルについて具体的に示すとともに、前記であげた通信コストの問題について検討する。

3.1 システムモデル

ドメインリーダーアルゴリズムは図1で示すように、システム上に複製を持ったサイトをグループに分割し、グループごとに複製を管理する。また、各グループをドメインとし、ドメイン内の1個のサイトをリーダーとする。リーダーは、ドメイン内のサイトデータの一貫性を管理し、他のリーダーとデータのやり取りを行う。リーダー間の同時実行制御は、プライマリという特殊なリーダーが行う。

また、ドメインリーダーでは、以下の方法を用いてデータの一致性と正当性を保証する。

a) プライマリコピー：リーダー間の処理を管理するためにプライマリコピー方式を用いる。このことから、1個のリーダーをプライマリとし、プライマリがリーダー間の同時実行制御を行う。

b) 2PC (2 Phase Commit) と 2PL (2 Phase Lock)¹¹⁾：データの正当性と一貫性を保証するために、リーダーどうしの通信を制御するために2PCを用いる。

c) 版番号：各サイトは、データの版番号を持っている。プライマリの版番号は、書き込みトランザクションごとにリーダーに転送される。これは、プライマリとリーダー間でデータが同一であることを確認するために利用する。リーダーも処理を実行する際、版番号を他のリーダーに、要求メッセージと一緒に配布する。これにより、全リーダーのデータが同一であることを、トランザクションごとに確認している。

なお、我々は以下の前提条件を考慮してアルゴリズムの動作や性能を検討する。

- ネットワークを論理的に完全グラフトポロジーとする。また、有線ネットワークを基本的なLANとし、全サイト間の通信速度を1Mbpsとする。

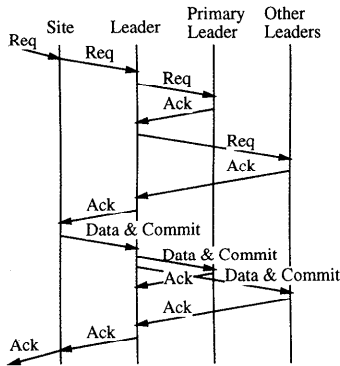


図2 サイトに到着した書き込み要求の処理 (DL)
Fig. 2 Procedure of write request arrived at a site.

- 全サイトは複製を持っている。また、1個のデータアイテムだけを考える。
- 全サイトは同じ処理能力、通信能力を持っている。また、同一なローカルデータベースシステムを持っている。

3.2 基本動作

サイトに書き込み要求が到着した場合、ドメインリーダーアルゴリズムでは、要求が到着したサイトと各ドメインのリーダーのデータだけを更新する。他のドメイン内のサイトデータは更新されない。これは、文献4)の動作とは異なるが、メッセージ数の削減のためにこのようなアルゴリズムとしている。これにより、サイトが読み込みを実行するときには、サイト自身とサイトが属するドメインのリーダーのみをアクセスすればよいので、読み込みの通信コストが小さくなる。

またサイトが書き込みを実行する際は、そのサイトとリーダーだけのデータを更新するので、書き込みの通信コストも小さくなる。

書き込みおよび読み込みの要求がサイトに到着した際のプロトコルの流れは次のようになる。

a) 書き込み要求が到着した場合

ここでは、リーダー以外のサイトに要求が到着した場合のみを考え、図2で示す。まず、サイトからトランザクションの要求がサイトの版番号を付けて発行される。次に、リーダーは、トランザクションの要求をプライマリに送信し、プライマリからのレディ(実行可)の応答を受け取ると、他のリーダーにトランザクションの要求を送る。全員から、レディの応答を受け取るとサイトにレディを返し、データをサイトから送信してもらい、それをリーダーに配送してコミット要求する。コミット応答が返ってきた段階で、サイトとプライマリにトランザクション終了応答を送る。

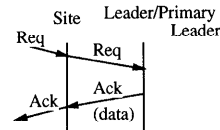


図3 サイトに到着した読み込み要求の処理 (DL)
Fig. 3 Procedure of read request arrived at a site.

b) 読み込み要求が到着した場合

サイトに要求が到着した場合、ドメインリーダーは図3に示すような動作を行う。まず、リーダーにトランザクション要求とともにサイトの版番号を送る。もし、リーダーが同じ番号のデータを持っていれば、同じ番号であることを返す。もし、リーダーの版番号が新しいものであれば、新しい版のデータをサイトに返送する。

4. ドメインリーダーアルゴリズムの拡張

この章では、ドメインリーダーアルゴリズムに、ある前提条件を加えることで、アルゴリズムの拡張を行う。前提条件としては、現在の計算機やネットワークは年々信頼性が高くなっており、コンピュータは簡単には故障せず、ネットワークに流したパケットメッセージの紛失は少ないと仮定する。このような条件では、1つのトランザクションを処理する際に、いくつかの確認メッセージを省略することができる。

4.1 拡張アルゴリズム

4.1.1 改良ドメインリーダー (DL-II)

図4で示すように、この改良アルゴリズムでは、リーダーは必ず存在し、問題なく動作していると仮定するので、リーダーの存在については調べない。また、基本的に他のドメインにあるサイトのデータの更新は行わない。また、全リーダーの応答を待つ状態を過半数だけにした。

過半数にした理由は、ネットワークが分断されたときでも、過半数の応答が返るサブシステムについてのみは処理を続行させるためである。

4.1.2 全サイト更新ドメインリーダー (DL-all)

ここでは、リーダー以外に、全サイトのデータを更新するアルゴリズムを検討する(図5)。アルゴリズムとしては、ほぼ前項で述べた改良アルゴリズムと同じであるが、各リーダーが全サイトを更新する動作を付け加えた。サイトのデータの読み込みは、リーダーに問い合わせることなく行われる。

4.1.3 シミュレーションによる性能評価

ここでは、旧ドメインリーダー (DL), 改良ドメインリーダー (DL-II), 全サイトを更新するドメインリーダー (DL-all) が読み込みと書き込みトランザクションを

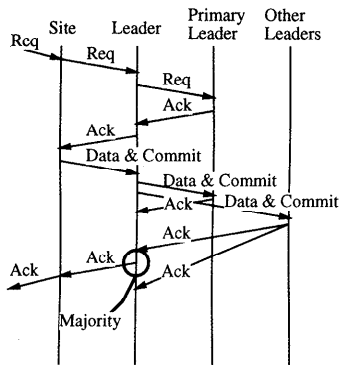


図4 サイトに到着した書き込み要求の処理 (DL-II)
Fig. 4 Procedure of write request arrived at a site.

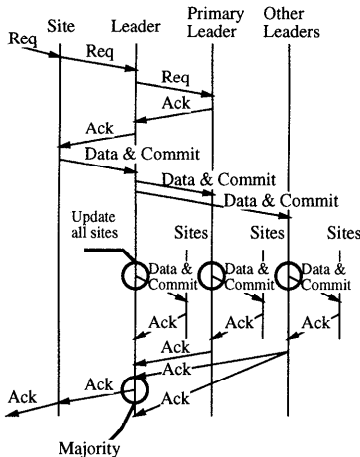
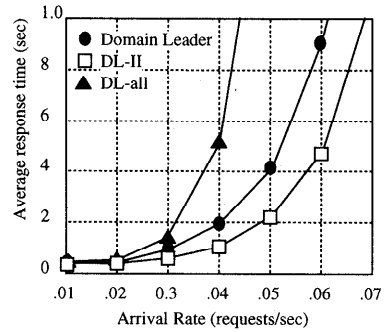


図5 サイトに到着した書き込み要求の処理 (DL-all)
Fig. 5 Procedure of write request arrived at a site.

処理するための平均応答時間を、要求の到着率を変化させて比較する。シミュレーションの条件を以下にまとめる。

- サイト数 100, リーダ数 10 (9 サイト/ドメイン)
- 通信速度は 1 Mbps
- サイトの処理時間はゼロとした
- 通信形態は非放送型
- 通常のメッセージは 20 byte, データは 1 Kbyte

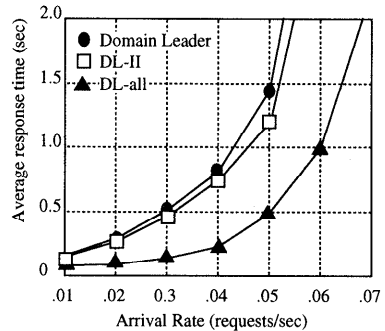
DL-II は、通常ドメインリーダー DL より 5% から 10% 応答時間を改善した (図 6, 7)。DL-all は、読み込み要求の発生率が少ない場合、応答時間は DL-II より悪いが (図 6, 7)、読み込みの到着率が高くなると応答時間が他のアルゴリズムより速くなる。これは、DL-all は読み込みの際、リーダーをアクセスする必要がないので、読み込み要求が多いシステムでは、応答時間が短くなるためと考えられる。



(site = 100, data = 1 Kbyte, read prob. = 10%, write prob. = 90%)

図6 平均応答時間の評価

Fig. 6 Evaluation of average response time.



(site = 100, data = 1 Kbyte, read prob. = 90%, write prob. = 10%)

図7 平均応答時間の評価

Fig. 7 Evaluation of average response time.

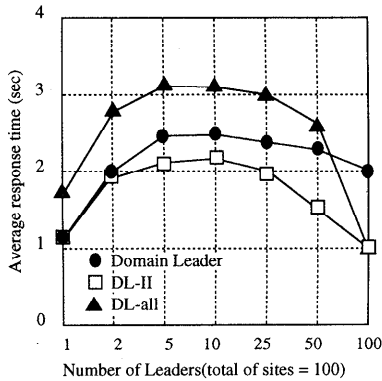
4.2 リーダとリーダー数の決めかた

ドメインリーダーアルゴリズムでは、サイトをいくつかのドメインに分割するが、それによってリーダー数やドメイン内のサイト数が変化し、通信コストが変化する。その最適な数を与えることは、システム設計において重要となる。

本節では、DL, DL-II, DL-all の平均応答時間をシミュレーションで評価し、その応答時間が一番低いリーダー数を最適なリーダー数とする。また、なんらかのサイト数 N の関数になっているかについて考察する。

図 8 は、サイト数が 100 で、要求の到着率が 0.001/秒で、読み込み要求の到着率が 10%、書き込み到着率が 90%、データのサイズが 1 Kbyte、サイトの処理時間が 0.1 秒の例を示す。

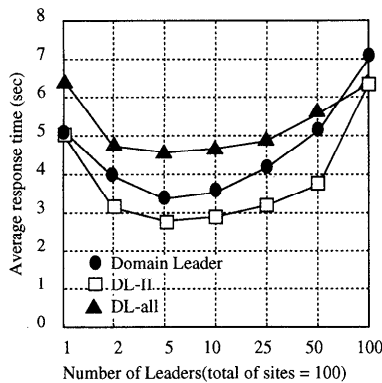
要求の到着率が小さい場合、要求が競合することがなく、応答時間はほぼリーダー間の通信とサイトの処理時間となる。この場合、リーダー数が小さいと書き込みのコストが減り、応答時間が短くなる。また、リーダー



(site = 100, size = 1 Kbyte, arriv. prob. = 0.001/s, read prob. = 10%, write prob. = 90%)

図8 平均応答時間の評価

Fig. 8 Evaluation of average response time.



(site = 100, size = 1 Kbyte, arriv. prob. = 0.01/s, read prob. = 10%, write prob. = 90%)

図9 平均応答時間の評価

Fig. 9 Evaluation of average response time.

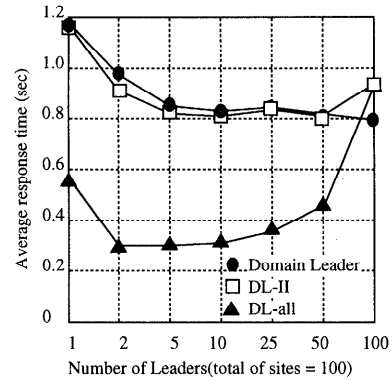
数が大きいと、書き込みのときのサイトとリーダー間の通信が減り、読み込みの際の通信コストが減少し、応答時間が短くなる。以上のことから、到着率が低い場合は、最適なリーダー数を求めることはできなかった。

次に、要求がキューで待たされるように到着率をあげて、応答時間を検討する。図9は、到着率を0.01/秒にし、他の条件は図8と同じときの各リーダー数の応答時間を比較した。結果として、読み込みと書き込みの到着率の割合が10%と90%、50%と50%の各条件で、データサイズに関係なくリーダー数が5のときに応答時間が最小となった。つまり、読み込み要求の到着率が50%までは、データサイズが変わっても、応答時間が最小となるリーダー数の割合は変わらなかった。さらに、サイト数が50個、200個、500個の各場合において、最適なリーダー数のサイト数に対する比率は変わらなかった。データサイズを1Kbyte、読み込みの到

表1 システム全体のサイト数と最適なリーダー数

Table 1 System site number and the best leader number.

サイト数	最適リーダー数	ドメイン内のサイト数
50	3	16
100	5	19
200	10	19
500	25	19



(site = 100, size = 1 Kbyte, arriv. prob. = 0.01/s, read prob. = 90%, write prob. = 10%)

図10 平均応答時間の評価

Fig. 10 Evaluation of average response time.

着率を10%、書き込み要求の到着する割合を90%と固定して、3つのアルゴリズムの応答時間が最小となったリーダー数を、表1にまとめる。

この表から、システム全体のサイト数を N とすると、最適なリーダー数 Le は、次のように求めることができる。

$$Le = N/20 \quad (1)$$

ただし、図10で示すように、到着した要求が90%読み込み要求の場合は、最適なリーダー数は現れず、式(1)は成立しなかった。今回のシミュレーションでは、すべての場合に成立するような最適なリーダー数を定式化することはできなかった。

4.3 他のアルゴリズムとの性能比較

プライマリコピー (PC)、定数合意 (QC)、階層型定数合意 (HQC)、三角格子 (TL)、そしてドメインリーダー (DL) の各方式について、サイトアクセス数を表2に整理した。定数合意では、読み込みと書き込みが同じ通信コストで処理できる状態とした。

ドメインリーダーのアルゴリズムとしてDL-IIを用いた。DL-allは書き込みの通信コストが高いため、検討しない。また、DL-IIが書き込みを実行するときアクセスを必要とするリーダー数として、4.2節の式(1)を使う。読み込み要求がリーダーに到着した場合、他のサイトをアクセスする必要がないので、1とする。サ

表2 各方式のサイトアクセス数
Table 2 Number of site access.

方式	PC		QC	HQC	TL		DL	
	r	w	r/w	r/w	r	w	r	w
サイト アクセス数	1	N	$\frac{N+1}{2}$	$N^{0.63}$	\sqrt{N}	$2\sqrt{N}-1$	2	$\frac{N}{20}+1$

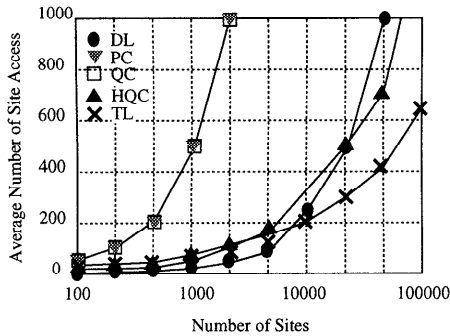


図11 サイトアクセス比較グラフ
Fig. 11 Comparison of site access number.

イトに到着した場合、リーダーをアクセスするので、2となる。

図11に、各アルゴリズムでのアクセスサイト数によって通信コストを評価した結果を示す。読み込みと書き込みの比率は平均で1:1として算出している。図で示すように、サイト数が約3000以下では、ドメインリーダーは他のアルゴリズムと同等なサイトアクセス数を示す。

5. 故障発生時の対処

本研究では、プロセスが正しく動作しないビザンティン故障は発生しないと仮定する。ここでは、システム障害やメディア障害によってプロセスが完全に停止する停止故障だけを考慮し、サイト故障とする。また、通信障害によって、サイトが他のサイト、リーダー、プライマリと通信できない状態もサイト故障とする。

5.1 ドメインリーダーの可用性

ドメインリーダーアルゴリズムでは、データの可用性を保証するために以下の方法を用いる。

a) ドメインテーブルとリーダーテーブル

ドメインテーブルは、ドメイン内の全サイトが持っているテーブルで、各サイトの優先度を記述したファイルである。このテーブルは、リーダーが壊れた場合、他のサイトを優先度順にリーダーを決めるために用いる。リーダーテーブルは、全サイトとリーダーが持っているテーブルで、各リーダーの優先度を記述したファイルである。これは、ドメインテーブルと同じように、プ

ライマリが壊れた場合に、新プライマリを決めるために使う。

b) 過半数サイト数

ドメイン内で複数のサイトが故障した場合、過半数のサイトグループと通信できるサイトについては通常に動作できるようにする。過半数のサイトと通信できないサイトは、孤立したものとして処理を中止する。

リーダーの場合も同様で、複数のリーダーが故障した場合、過半数のリーダーと通信できるリーダーグループだけがトランザクション処理をすることができる。

c) ログファイル

故障が発生したあとの回復制御のときに使う。このファイルには、更新する前のデータ、更新されたデータなどの情報がある。また回復後、処理中であったトランザクションが完了したか、アボートされたか、このまま続行できるか判断するために必要なデータが記述されている。

5.2 故障対処のメッセージ数の比較

ここでは、故障を検出した後に、故障を対処する際、プライマリコピー方式、三角格子方式、およびドメインリーダーではどのくらいのメッセージ数が必要かを検討する。

5.2.1 プライマリコピー方式

故障を検出したサイトは、他のサイトにサイトが故障しているというメッセージをマルチキャストして知らせる。他のサイトから返事が到着した場合、サイトは、他のサイトに全サイトが1つのサイト故障が認定されたというメッセージをマルチキャストする。以上のように3つのメッセージが必要である。このことから、Nをシステム全体のサイト数とすると、故障に対処するために、以下のメッセージ数が必要である。

$$Mp = 3(N - 2), \quad N > 1 \tag{2}$$

5.2.2 三角格子

三角格子では、論理的にサイトが格子状に接続されていると考えている。故障を検出したサイトは、その故障をしたサイトが含まれる行と列の各サイト、および故障の際に使われる迂回サイトに故障を通知する。その確認の応答が通知した全サイトから戻ったとき、迂回路を使った新しい行と列のアクセスサイトが通知

した全サイトに配送される。

このメッセージ数は、行と列および迂回サイトを合わせたがサイト数が $2N$ であるから、

$$Mt = 6N, \quad N > 1 \quad (3)$$

となる。

5.2.3 ドメインリーダー

ドメインリーダーでは、サイトが壊れた場合、ドメイン内で故障対処の処理を行うため、リーダー数を L とすると、ドメイン内のサイト数は N/L となる。結局、式 (2) から、

$$Md = 3(N/L - 2), \quad N > L, \quad L > 1 \quad (4)$$

となる。また、全リーダーにサイト故障を知らせるので、

$$M_{DL} = 3(N/L - 2) + 3(L - 1), \quad N > L, \quad L > 1 \quad (5)$$

となる。

以上のことから、ドメインリーダーの故障の対処に必要なメッセージ M_{DL} はプライマリコピーのメッセージ数 M_p より小さいが、三角格子のメッセージ数 M_t より大きい。ただし、 L が $N^{0.5}$ に近い値であれば、ドメインリーダーのメッセージ数は三角格子のメッセージ数に近づくことになる。

6. モバイル環境への適用

6.1 モバイル環境の特徴

現在、計算機技術と通信技術の進歩により、携帯端末を用いてワイヤレス通信や移動通信するモバイル環境が増えつつある^{1),2)}。携帯端末は、ノートパソコンのような小型計算機であって、PHSや携帯電話を用いることにより他のコンピュータと通信できる。しかし、通常の有線で接続されているコンピュータ間の通信に比べて、通信速度の面で2桁程度の差があると考えられる。

このような環境で、比較的大規模な複製に対して通常の複製管理方式を使おうとすると、通信コストがかかり性能が悪くなる。プライマリコピー方式では、読み込みのコストが低く、読み込みがほぼ100%のシステムにおいては利用できるが、そのような応用は少ない。ドメインリーダー方式は、読み込みのコストがプライマリコピーと同程度で小さいうえ、書き込みのコストはプライマリコピーより少ない。また、システム構成にかかる手間も、ドメインというグループ分けだけであるため、階層型定数合意や三角格子アルゴリズムより簡単である。したがって、応用は広いと考えられる。また、ドメインリーダーのリーダーどうしの通信と、サイトとリーダー間の通信頻度が異なる特性を利用することで、モバイル環境に最適な構成とすることもで

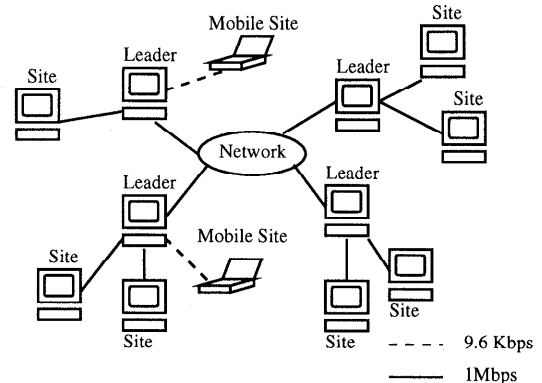


図12 モバイル環境
Fig. 12 Mobile environment.

きる。以下に、ドメインリーダーアルゴリズムをモバイル環境に適用した場合について評価する。

6.2 対象システムモデル

我々は、図12のようにネットワーク上に複数のモバイル端末が存在し、データベースをアクセスする環境を仮定し、ドメインリーダーアルゴリズムを評価する。モバイル端末はPHSか携帯電話を用いて9.6 Kbpsの通信速度、サイトは1 Mbpsの通信速度でデータの送受信を行う。また、ネットワークの信頼性が高くサイトの故障が発生しないとする。

6.3 性能評価

ここでは、表2のサイトアクセス数比較表に基づいて、DL、QC、HQC、TLの通信コストを比較する。また、ドメインリーダーのリーダー数を、式(1)によって算出する。

6.3.1 1個のモバイルサイトの通信コスト

まずはじめに、10%のモバイルサイトが含まれる環境において、1個のモバイルサイトが自分のデータを更新したい場合どのくらいの通信コストが必要かを示す。簡単のため、メッセージのデータサイズは平均で1KBとし、ユニキャスト通信を使うと仮定している。また、DLを除く方式では、モバイルサイトが9.6 Kbpsの通信リンクを使って2PCに基づく通信を行う(メッセージが2往復する)。DLでは、モバイルサイトからリーダーまでは9.6 Kbpsのリンクを使う。また、リーダーがモバイルサイトであればリーダーもすべて9.6 Kbpsのリンクとなり、リーダーが固定サイトで1 Mbpsのリンクを使える場合は、そのリンクを使う。そのようにして、通信コストを計算した結果を図13に示す。

図13では、サイト数が多くても(10000サイト程度)ドメインリーダーは他のアルゴリズムより低い通

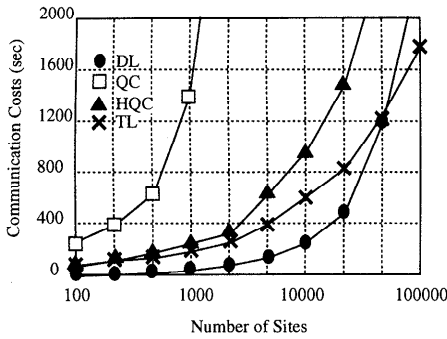


図 13 1 個のモバイルサイトの書き込み処理時間
Fig. 13 Write processing time of 1 mobile site.

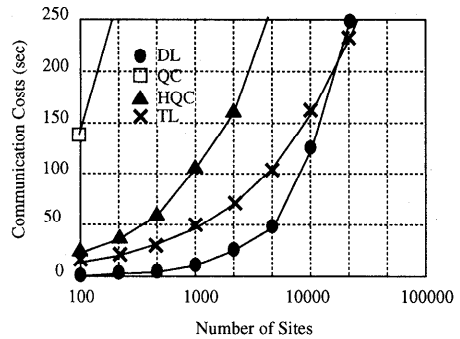


図 15 モバイルサイトがシステム全体の 50%のときの通信コスト
Fig. 15 Communication cost (mobile site is 50% of total site).

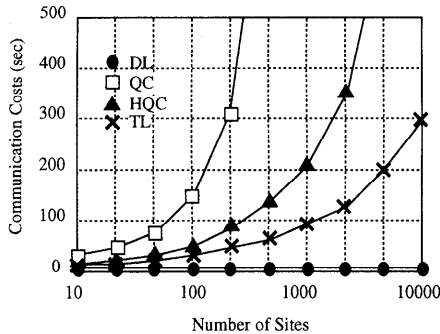


図 14 1 個のモバイルサイトの読み込み処理時間
Fig. 14 Read processing time of 1 mobile site.

信コストとなった。読み込みの際は (図 14), ドメインリーダーでは、リーダーだけアクセスすればよいので、低い通信コストとなる。

6.3.2 システム全体の通信コスト

ここでは、システム全体のサイト数が 50%モバイルサイトである場合の、平均の通信コストを示す (図 15)。ただし、書き込みと読み込みの比は 1:9 としてある。図に基づいて説明すると、モバイルサイトが多い環境では、他のアルゴリズムは通信コストが増加し、効率が悪くなるが、ドメインリーダーアルゴリズムは、比較的低い通信コストを維持している。また、逆にモバイルサイトが少ない環境 (均一環境) では、ドメインリーダーは、サイト数が多い場合、そして書き込み比率が大きくなると、従来の方式と比べて通信コストが大きくなるのが分かった。

7. おわりに

本論文では、大規模な分散システムの複製管理用に提案されたドメインリーダーアルゴリズムの性質について、詳細に評価した。また、特定の環境に適した拡張を提案し同時に評価した。

ドメインリーダーを改善した DL-II は、ドメインリー

ダに対して読み込みと書き込みを処理する際にかかる応答時間を 5%から 10%短縮した。読み込み要求が多い場合 DL-all は、ドメインリーダーと DL-II より短い応答時間で要求を処理した。ドメインリーダーアルゴリズムは、要求の到着率が比較的高い領域では、全体のサイト数に応じた適切なリーダー数があることが明らかとなった。

サイトが故障した場合、ドメインリーダーテーブル、リーダーテーブル、過半数サイト数、ログファイルを用いて、可用性を保証した。これにより、故障に対応するために必要とするメッセージ数はプライマリコピーより少ないことを示した。

モバイル環境を想定した通信コストの比較により、従来の均質の環境を想定した複製管理アルゴリズムが必ずしも性能的に優れておらず、ドメインリーダー方式が不均質ネットワークに適していることが明らかとなった。

また、一般的な最適リーダー数の定式化は今回のシミュレーションからは求められなかった。最適なリーダー数が、サイト数を N として N より小さい値となれば、サイトアクセス数が従来の方式より良くなると考えられる。今後、より様々な環境でシミュレーションすることで見つけていきたい。

参考文献

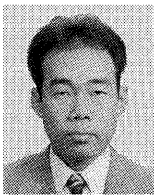
- 1) 水野忠則, 田窪昭夫: モバイルコンピューティング, 情報処理, Vol.36, No.9, pp.822-826 (1995).
- 2) 水野忠則: モバイルコンピューティングの現在と未来, 情報シンポジウム論文集, Vol.95, No.10, pp.9-16 (1995).
- 3) 成田藤智, 村瀬 亨, 塚本昌彦, 西尾章治郎: 移動型データベースのためのビューの実現における ECA ルール配送機構について, 情報処理学会研

究報告, Vol.96, No.MBL-2, pp.7-12 (1996).

- 4) 宮西洋太郎, 中村健二, 佐藤文明, 水野忠則:分散システムにおけるデータの複製管理方式, 情報処理学会論文誌, Vol.37, No.5, pp.830-839 (1996).
- 5) Gifford, D.K.: Weighted voting for replicated data, *Proc. 7th ACM SIGOPS Symp. Oper. Syst. Princip.*, CA, Dec. 10-12, pp.150-159 (1979).
- 6) Kumar, A.: Hierarchical quorum consensus: A New Algorithm for Managing Replicated Data, *IEEE Trans. Comput.*, Vol.40, No.9, pp.996-1004 (1991).
- 7) Wu, C. and Belford, G.: The Triangular Lattice Protocol: A Highly Fault Tolerant and Highly Efficient Protocol for Replicated Data, *Proc. 11th Symp. on Rel. Distrib. Syst.*, pp.66-73 (1992).
- 8) 中村健二, 宮西洋太郎, 佐藤文明, 水野忠則:ドメインリーダーによる複製管理プロトコルの提案, 情報処理学会全国大会論文集, pp.91-92 (1995).
- 9) 中村健二, 宮西洋太郎, 佐藤文明, 水野忠則: Replica Control Protocol Based on Domain Leader Concept, 情報処理学会研究報告, Vol.95, No.115, pp.93-98 (1995).
- 10) 中村健二, 宮西洋太郎, 佐藤文明, 水野忠則:ドメインリーダーに基づく複製管理方式のモバイル環境適用への評価, 情報処理学会研究報告, Vol.96, No.288, pp.43-48 (1996).
- 11) 喜連川優 (監修), 岩宮好宏ほか (著):分散トランザクション処理, リックテレコム (1994).

(平成9年5月7日受付)

(平成9年12月1日採録)



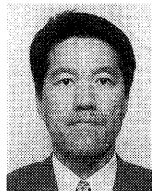
中村 健二 (正会員)

昭和45年生。平成6年ブラジル国立パラ総合大学情報技術学科卒業。平成6年静岡大学工学部研究生。平成9年同大学大学院工学研究科修士課程情報知識工学専攻修了。同年日本ヒューレット・パカード(株)入社。現在に至る。主な研究分野は、分散データベース、モバイルコンピューティング。



宮西洋太郎 (正会員)

昭和18年生。昭和43年神戸大学大学院工学研究科電気工学専攻修了。昭和43年三菱電機(株)入社。主に工業分野における各種応用システムの構築にSEおよび管理者として従事した。在職中、平成9年静岡大学大学院電子科学研究科電子応用工学専攻修了、工学博士。現在、同社情報システム製作所に勤務、分散システムの性能評価、分散DB複製配置、企業間情報共有方式(建設分野へのCALS技術適用)などの研究および業界活動に従事。共訳として「分散オペレーティングシステム」(プレントイスホール)、共著として「CALSの実践」(共立出版)がある。専修大学経営学部非常勤講師。計測自動制御学会、システム制御情報学会各会員。



佐藤 文明 (正会員)

昭和37年生。昭和61年東北大学大学院工学研究科電気及通信工学専攻博士前期課程修了。同年三菱電機(株)入社。通信ソフトウェアの研究開発に従事。平成7年1月より静岡大学工学部助教授。現在、静岡大学情報学部助教授。工学博士。通信ソフトウェア、形式言語、分散処理システムに関する研究に興味を持つ。電子情報通信学会、IEEE Computer Society 各会員。



水野 忠則 (正会員)

昭和20年生。昭和43年名古屋工業大学経営工学科卒業。同年三菱電機(株)入社。平成5年静岡大学工学部情報知識工学科教授。現情報学部情報科学科教授。工学博士。情報ネットワーク、プロトコル工学、モバイルコンピューティングに関する研究に従事。著書としては、「プロトコル言語」(カットシステム)、「MAP/TOPと生産システム」(オーム社)、「分散システム入門」(近代科学社)、「分散システム—コンセプトとデザイン」(電気書院)などがある。電子情報通信学会、IEEE 各会員。