

音声認識機能付き VoIP システム

## 評価実験報告書

## 目次

1.VoIP システム.....	3
1.1 通信品質(遅延)の評価実験.....	5
1.2 ユーザビリティの評価.....	8
1.3 経済性の評価.....	9
2.音声認識/合成の評価実験.....	15
2.1 音声認識率の評価実験.....	15
2.2 ユーザビリティの評価実験.....	16
2.3 アクセシビリティの評価実験.....	18
3.非同期システム.....	20
3.1 アベイラビリティの評価実験.....	20
3.2 ユーザビリティの評価実験.....	22
3.3 スケーラビリティの評価実験.....	27
4. 結言.....	29

[実験環境に関して]

実施計画書中、評価内容に関して

VoIP システムについて、50 人以上(1 基地局同時 10 人以上)の被験者により、以下の実験及び評価を行う。

- ・無線環境下での VoIP 電話の通話品質評価: 目標、300 ミリ秒以下の遅延
- ・ユーザビリティ (ISO13407 の基準等) の検証
- ・利用頻度、利用エリアの分析と経済性の評価

PDA 用音声認識について、50 人以上(1 基地局同時 10 人以上)の被験者により、以下の実験及び評価を行う。

- ・音声認識率の測定評価
  - ・音声認識及び音声による情報検索の効率測定評価: 目標、音声認識 1 フレーズあたり 2 秒以内、検索 10 秒以内
- ・ユーザビリティ (ISO13407 の基準等) の検証
- ・アクセシビリティの評価

音声サーバシステムソフトウェアについて、50 人以上(1 基地局同時 10 人以上)の被験者により、以下の実験及び評価を行う。

- ・アベイラビリティの検証: 目標、99.9%以上(年間 8.5 時間のダウンタイム)
  - ・ユーザビリティ (ISO13407 の基準等) の検証
  - ・スケーラビリティの検証 : 目標、高負荷時 2 秒以内の応答時間
- と記載されているが、同環境下で試験を行ったのは、

[50 人以上(1 基地局同時 10 人以上)の被験者の環境で実験すべき項目]

VOIP :

無線環境下での VoIP 電話の通話品質評価: 目標、300 ミリ秒以下の遅延

音声(非同期)サーバシステム :

アベイラビリティの検証: 目標、99.9%以上(年間 8.5 時間のダウンタイム)

スケーラビリティの検証 : 目標、高負荷時 2 秒以内の応答時間

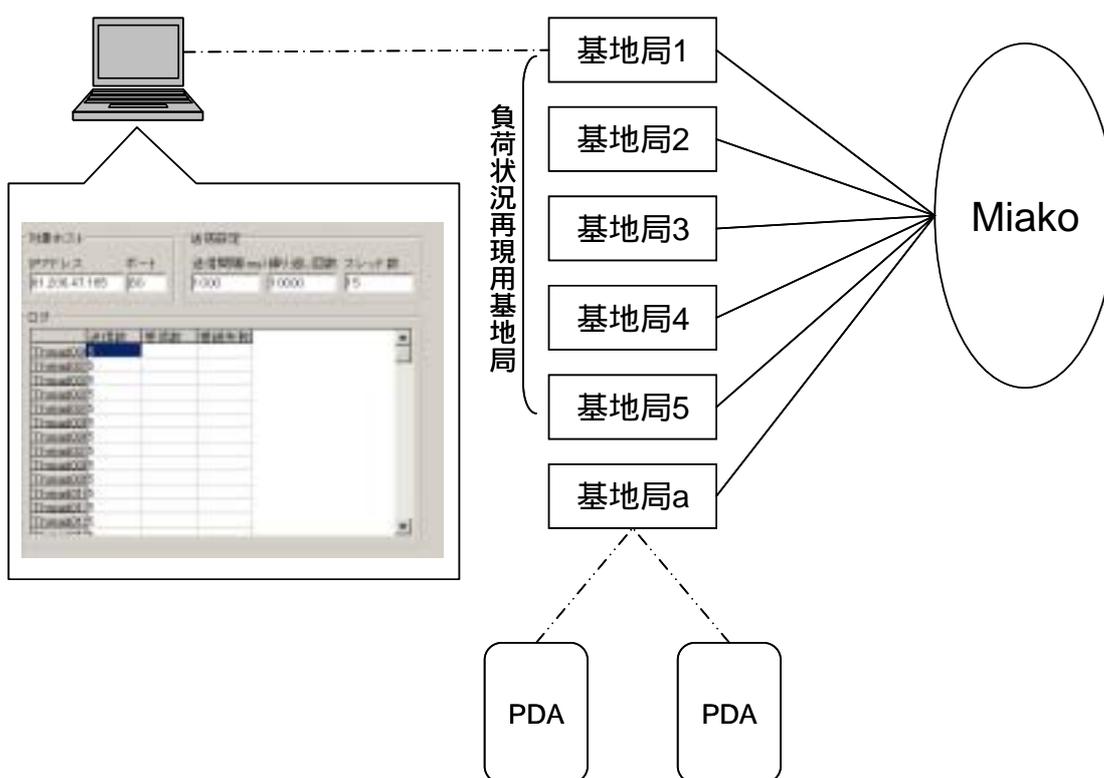
の 3 項目である。これは上記項目がネットワークを用いて、高トラフィック下での性能評価を行うことが目的であるのに対し、その他項目はネットワークを用いない機能であり、条件的に意味をなさない為である。

また、50 人以上(1 基地局同時 10 人以上)のトラフィック環境は、時間的人的に再現が困難であるため、以下の方法を用いて 50 人以上(1 基地局同時 10 人以上)の被験者環境を構築した。

- ・無線基地 5 局に各 1 台 PC を設置
- ・各 PC に負荷テストプログラムを入れ、1s 毎に HTTP リクエストを 15(人分)スレッド発信

これにより仮想的に 1 基地局 10 名以上が一般的なネット接続が行われた(1s 毎にリクエストを発生しているため、現実の利用環境より高負荷と考えられる)環境を構築した。

50 人以上(1 基地局同時 10 人以上)環境の構築



負荷テストプログラムは以下のソフトウェアを使用した。

<http://homepage3.nifty.com/fireballz/freeware/shark/readme.txt>

なお、上記で想定した人数(50 人以上 1 基地局同時 10 人以上)は、全て当該システムを利用していることは想定しておらず、あくまで通常のインターネット利用(WWW,Mail etc)を想定したものであり、通常の利用形態と当該システム間での干渉の有無を検証するものである。

なぜなら、当該実験は当該システムソフトウェアの性能を検証するものであり、当該システムで用いられた無線 LAN インフラ miako.net の性能を検証するわけではない。

高負荷状況におけるネットワークトラフィックの問題は、インフラで用いるバックボーン回線の帯域に左右されるものであり、ソフトウェア(含むシステム)の評価とは別である。

## 1. VoIP システム

### 1.1 通信品質(遅延)の評価実験

#### [実験概要]

PDA を用いた P2P VoIP システムでの音声通話に関し、遅延状況を検証することで、実用度の評価を行った。

実験は以下の条件で行った。

#### 通信インフラ

- ・miako-net(<http://www.miako.net/>)環境を利用
- ・1 基地局 10 人以内の接続状況下
- ・miako-net 上で 50 人以上の利用状況下(または同程度の負荷時)で測定

#### 通話端末

- ・PocketPC2002 を搭載した PDA
- ・802.11 に対応した無線 LAN カード

#### [計測方法]

- ・上記通信インフラを用いて 2 局間で通話
- ・通話は 1:1 で行う
- ・通話遅延を 300ms 以下に設定し、通話が良好か測定
- ・全 20 人以上の被験者を用いて実験

#### [目標値]

300ms 以下の音声到達遅延

#### [実施内容]

実証実験は miako-net 環境下の複数の基地局を用いて、アプリケーションの遅延パラメータを 300ms 以下に設定し、以下の日程で行った。

実験実施年月日	被験者甲実験場所	被験者乙実験場所
2003年 2月 3日	京都市中京区蛸薬師通烏丸 西入ル ヒライビル 3F 株式会社 ネットイン京都	京都市中京区蛸薬師通烏丸 西入ル ヒライビル屋上 株式会社 ネットイン京都
2003年 2月 4日	下京区東洞院通七条下ル東 小路町676 - 13 1階ロビー ぱるるプラザ京都	左京区宝ヶ池財団法人国立京 都国際会館 1階 国立京都国際会館
2003年 2月 6日	中京区小川通六角下ル元本	京都市中京区蛸薬師通烏丸

	能寺町 SCCJ 事務局	西入ル ヒライビル屋上 株式会社 ネットイン京都
2003年 2月 10日	京都市東山区川端通四条上 ル常盤町 井筒八ッ橋祇園店	北区小山北上総町 16 北大路 商店街
2003年 2月 11日	下京区烏丸通七条下ル 京都 タワービル1階 国際観光振興会京都案内所	中京区堀川通二条城前 京都 全日空ホテル 1階ロビー
2003年 2月 12日	下京区東洞院通七条下ル東 小路町676-13 ぱるるプラ ザ京都1階ロビー	京都市東山区下河原町 高台寺美術館
2003年 2月 13日	左京区粟田口烏居町2番地の 1 京都市国際交流協会ホー ル	西京区大原野西境谷
2003年 2月 14日	京都市中京区蛸薬師通烏丸 西入ル ヒライビル3F 株式会社 ネットイン京都	京都市西京区大原野西境谷
2003年 2月 17日	京都市中京区堀川通二条城 前 京都全日空ホテル 1階ロ ビー	京都市北区小山北上総町 北大路商店街
2003年 2月 19日	中京区姉小路通東洞院東入 笹屋町	上京区烏丸通り一条下ル龍前 町 KBS カルチャー
2003年 2月 25日	中京区三条通御幸町角 1928 ビル地下1階 カフェアンデパ ンダン	中京区姉小路通東洞院東入 笹屋町
2003年 2月 28日	中京区烏丸蛸薬師株式会社 ネットイン京都	中京区堀川通二条城前京都 全日空ホテル
2003年 3月 3日	中京区烏丸蛸薬師株式会社 ネットイン京都	上京区烏丸通り一条下ル龍前 町 KBS カルチャー
2003年 3月 4日	京都市東山区祇園町舞風館 別館花楽	京都市中京区蛸薬師通烏丸 西入ル ヒライビル 3F 株式会 社 ネットイン京都
2003年 3月 5日	中京区蛸薬師通烏丸西入ル ヒライビル3F 株式会社 ネット イン京都	京都市中京区蛸薬師通烏丸 西入ル ヒライビル屋上株式会 社 ネットイン京都

2003年 3月 6日	京都市中京区堀川通二条城 前 京都全日空ホテル 1階ロ ビー	京都市北区紫野花ノ坊町 11 番地 京都ライトハウス1階
2003年 3月 10日	京都市北区小山北上総町 16 北大路商店街	京都市東山区祇園町舞風館 別館花楽
2003年 3月 12日	京都市左京区宝ヶ池財団法人 国立京都国際会館 1階	伏見区塩屋町 239-1 龍馬館
2003年 3月 19日	京都市下京区烏丸通七条下 ル 京都タワービル1階 国際観光振興会京都案内所	京都市中京区蛸薬師通烏丸 西入ル ヒライビル 3F 株式会 社 ネットイン京都
2003年 3月 20日	京都市中京区蛸薬師通烏丸 西入ル ヒライビル 3F 株式会 社 ネットイン京都	京都市中京区蛸薬師通烏丸 西入ル ヒライビル屋上

実験は被験者甲及び被験者乙が交互に3分以上の通話を行う。またこの際上記に示した負荷環境により仮想的に50人以上(1基地局10人以上)のネットワークトラフィックを再現している。VoIPアプリケーションは遅延パラメータに基づきパケットを送信するため、300ms以上の遅延が発生した場合、パケット落ち(データの送受信もれ)が発生し、通話が途切れ途切れになる。したがって、上記設定で安定した通話が行えれば300ms以下の遅延が達成できたことになる。しかしながら、遅延発生の原因に関してはアプリケーション以外に、ネットワークトラフィックによる場合も考えられるが、当該目標値はあくまでソフトウェアの性能としての遅延をあげたものであり、昨今各地で発生しているウィルス被害によるトラフィック低減のために発生した遅延に関しては、これを含まない。実験は上位ネットワークのトラフィックに影響されない miako-net 内だけの通信で行った。

#### [実施結果]

通話遅延に関しては全く問題がなかったが、交通量の多い箇所、アンテナと被験者の位置関係等が原因で、通話品質(声質)の低下、ノイズの混入が見られる場合があった。

これは無線の電波状況による問題であり、通話品質が低下した場合も遅延は発生しなかった為、実証に関しては問題ない。

300ms以上の遅延が発生しなかったことで、通話システムとしては十分実用的であると言える。当該システムでは、遅延発生を避けるため、システムに余計な負荷をかけないよう、通話時の音声圧縮を最低限しか行っていない。

今後はアプリケーションロジックの最適化などにより、より高度な音声圧縮を実装することにより、通話品質の向上に取り組むことが課題となる。

## 1.2 ユーザビリティの評価

### [実験概要]

VoIP 音声通話におけるユーザーインターフェースに関して、ISO13407(インタラクティブ・システムのための人間中心主義的設計)で定義されるヒューマンインターフェースデザインに準拠して設計されたかを評価した。

### [計測方法]

インスペクション法による評価

被験者 5 名によるヒューリスティック評価と問題点の解析

### [目標値]

ISO13407 で定義されるヒューマンインターフェースデザインに準拠し、かつ簡素で使い勝手の良いインターフェースデザインがなされていること

### [実施内容]

ユーザビリティの評価に関しては明示的に数値により計測するような評価方法がないため、インスペクション法(研究者、開発者、デザイナー等が経験に基づいて問題点を見つける手法)に基づく評価を行った。

評価は 5 名の被験者(ヒューリスティック評価における標準的な人数：当該システム SE1 名、PG1 名、デザイナー1 名、音声関連研究者 1 名、視覚障害者 1 名)により、各位がニールセンのユーザビリティ 10 原則に基づくヒューリスティック評価項目に関しての評価を行い、最終的に各位の評価をまとめ、問題点の解析等を行い、評価結果とした。

### [実施結果]

以下、ヒューリスティック評価項目に基づく評価内容を記載。

1.システム状態の視認性を高める	システム状態に関しては、視覚のみの視認性というだけでなく、常に状況やステータスを音声ガイダンスすることにより、ユニバーサルデザインとしての視認性を確保している。
2.実環境に合ったシステムを構築する	利用環境、利用者層を想定し、ガイダンス音声等の統一化を行っている。
3.ユーザにコントロールの主導権と自由度を与える	メニュー選択箇所などをループさせる(メニュー最後で行き止まりにならない)ことにより、選択の自由度をあげている。
4.一貫性と標準化を保持する	操作に用いられるハードウェアキーを他機能と統一することにより、オペレーションの一貫性を保っている。

5.エラーの発生を事前に防止する	通信に関して予期しないエラー(基地局のダウン)等が発生した場合、(すなわち自システムではなく、外部要因におけるエラーの場合)が発生した場合の防止手段がない。
6.記憶しなくても、見ればわかるようなデザインを行う	これまでのVoIP電話のようにIPアドレスを入力して電話というオペレーションではなく、携帯電話と同等の機能、オペレーションであり理解しやすい。
7.柔軟性と効率性を持たせる	通話を行う方法が1つではなく複数(音声認識、電話帳検索、着発信履歴)あることで、ある機能が不可だった場合の代替手段(例えば音声認識ができなかった場合に、電話帳検索する等)が用意されており、柔軟性が高い。
8.最小限で美しいデザインを施す	晴眼者からは、昨今のWebやアプリケーションデザインに比べると地味な印象はあるが、実用的には問題ない。
9.ユーザによるエラー認識、診断、回復をサポートする	5と同等で外部要因におけるエラー発生時の診断と回復は難しい。
10.ヘルプとマニュアルを用意する	アプリケーションレベルで音声ガイダンスが用意されており、そのままヘルプとして利用できる。

上記評価結果により、外部要因におけるエラー時の復旧に若干課題が残るが、正常時のユーザビリティには優れているという結果になった。

エラー時の診断回復は重要な要素であるが、外部要因におけるエラー(例えば上位通信インフラの不具合など)に関しては、現状では外部要因の復旧を待つしか手法がなく、アプリケーション側で復旧は不可能だがエラー認識はできており、ガイダンスもなされていることから、当該システムとしては目標値に達しているとの結論となった。

### 1.3 経済性の評価

1.1、1.2 で得られた評価結果及び、各無線基地局のアクセスログより得られた利用頻度などのデータを元に、当該システムの経済性に関して、以下の観点より評価を行った。

#### 既存技術と比較した技術的優位性の評価

既存技術と比較した場合、当該システムでは以下の技術的優位性があげられる。

- A.既存電話網と比較した場合の通話コスト低減
- B.無線 LAN と PDA を用いた移動体による VoIP 通話
- C.P2P 通信によるゲートウェイ設備のコスト低減
- D.オーブンプロトコルを用いた互換性

## E.無線 LAN 暗号化技術によるセキュリティ強化

### A.既存電話網と比較した場合の通話コスト

VoIP 技術を用いた音声通話が、既存電話網に比べ通話コストが低減されることは周知の事実であり、当該システムに限ることではない。特に遠距離通話になるほどそれが顕著になる。しかし、当該システムの場合、miako.net インフラを用いることにより、各アクセスポイントに必要な通話料を各アクセスポイント設置者が負担することで、他 VoIP サービスに比べ利用者の初期投資負担の提言、既存 VoIP 技術以上の通話コストの低減が可能である。

### B.無線 LAN と PDA を用いた移動体による VoIP 通話

現在、市販されている VoIP 技術は固定電話による通話を行う製品が主流であり、移動体による VoIP 通話は実用化されていない。市場傾向を見ると固定電話にかわり携帯電話が主流となっていることが明らかであり、これは今後 VoIP 市場に関しても同傾向になると見られる。当該システムでは無線 LAN と PDA を用いた移動体による VoIP 音声通話、また移動体通信に必要な基地局範囲切り替え時のハンドオーバーを実現しており、実用的な移動体による VoIP 通話を行える。

### C.P2P 通信によるゲートウェイ設備のコスト低減

既存の VoIP システムでは認証、互いの端末の名前解決等に用いるゲートウェイサーバシステムが必要であり、その為の設備投資、維持費が必要となる。

当該システムでは互いの固定 IP アドレスにより Peer To Peer による通話を行うため、ゲートウェイサーバが不要である。今後 IPv6 による現在の IP アドレス不足問題が解決された場合、当該システム型 VoIP システムが主流になると予想される。

### D.オーブンプロトコルを用いた互換性

当該システムは音声プロトコルに NOTASIP 方式を用いている。NOTASIP 方式は音声通話に特化したプロトコルであり、映像通信やデータ共有も視野に入れたプロトコルに比べ、シンプルであり C で記述したように特別なサーバシステムを必要としない。

また NOTASIP はオープンなプロトコルである。既存 VoIP システムでは各自が個別のプロトコルや通話方式を用いているため、キャリア毎の互換性がない。NOTASIP は既に様々な製品で実用化されていることから、異デバイス間での通話の互換性も保たれている。

## E.無線 LAN 暗号化技術によるセキュリティ強化

当該システムの通信インフラである miako.net は通信レベルでのデータ暗号化が行われており、盗聴などのセキュリティ対策にも優れている。

VoIP サービスの市場規模は 2007 年に 20 兆円以上(全世界)になると予想されており、また、2005

年には VoIP サービスの国内マーケットシェアは固定電話にならび 50%になると予想されている。  
(ともに IDC 調査)

特にアジア地区におけるブロードバンド普及に伴う VoIP 市場の拡大が予想されており、それに伴いモバイル機器を用いた VoIP サービス市場も急速に拡大すると予測される。

上記のように既に次世代音声通話網としてニッチな市場ではなくなった VoIP サービスであるが、爆発的な普及にはいくつかの問題が残っている。

- A. VoIP 技術の標準化
- B. 相互接続(互換)性
- C. セキュリティ対策
- D. エリア拡大までの対応
- E. 付加価値

#### A. VoIP 技術の標準化

VoIP サービスの普及に技術の標準化は欠かせない要因である。標準化は後述の相互接続性にも大きな影響を及ぼすため、早い段階での標準化が求められている。

現在、VoIP サービスにおけるプロトコル(通信手法)として

- ・H323
- ・SIP (Session Initiation Protocol)
- ・NOTASIP (Nothing Other Than A Simple Internet Phone)

が使われているが、当該システムでは NOTASIP 方式を用いている。これは他プロトコルが音声通話以外の映像通信やデータ共有を考慮して複雑化していったのに対し、音声通話のみに特化したシンプルなプロトコルであり、また P2P によりデータ送信を行うことから、ネットワークのトラフィックを圧迫しないためである。

現在 IPv6 や光回線によるネットワーク網の整備が急ピッチで行われているが、NOTASIP は現状及び上記インフラの整備後にもネットワーク上で世界規模の電話サービスを提供できる唯一の技術である。

当該システム開発で得られたノウハウ、及び技術をオープンにしていくことで、VoIP 技術の標準化に寄与する必要があるだろう。

#### B. 相互接続(互換)性

上記 NOTASIP 方式を用いることにより、NOTASIP 方式を用いたデバイス間での相互接続性を確保することが可能となった。当該システムでは PDA 間のみならず、PDA<->PC 間、PDA<->一般電話(携帯電話)との接続を行い、相互接続性の検証を行い、完全に互換性があることを確認した。

#### C. セキュリティ対策

当該実験インフラである miako.net では VPN による個人認証と通信データの暗号化を行っているため、無線電波の盗聴などセキュリティ対策は万全である。また当該システムの稼動する PDA(PocketPC2002)には VPN が標準搭載されているため、市販の無線 LAN カードがあれば特別なソフトウェアや機材を用いることなく安全な音声通話を行うことが可能である。

#### D. エリア拡大までの対応

ホットスポット(無線基地局)エリアの拡大は、当該システムの通話範囲と直接関係するため重要であるが、コスト面の問題から急速なエリア拡大は困難である。

既存の固定電話型 VoIP システムではこのような問題は発生しないが、移動体通信を行う場合、エリア外での対応をどうするのか考える必要がある。

当該システムでは非同期メッセージシステムを用いてこの問題に対処している。

現在、携帯電話キャリアにおけるデータ通信量は音声通信量を超えていると言われている。これは昨今の電子メール、電子掲示板等の非同期メッセージシステムに依るところが大きい。

時間に束縛された現代人にとって、電話のように自分や相手の時間を拘束しない非同期メッセージシステムは理想的なコミュニケーション手法と言える。

しかしながら、匿名性の高い文字情報でのコミュニケーションでは、なりすまし行為、迷惑メールなどの弊害も伴う。

そこで当該システムでは文字情報より信頼性の高い肉声によって非同期コミュニケーションを行うメッセージングシステムを用いた。

これにより相手がエリア外であっても、メッセージングシステム上に音声メッセージを残すことができる。また非同期であることから時間を拘束しないという電子メール的メリットも生まれる。

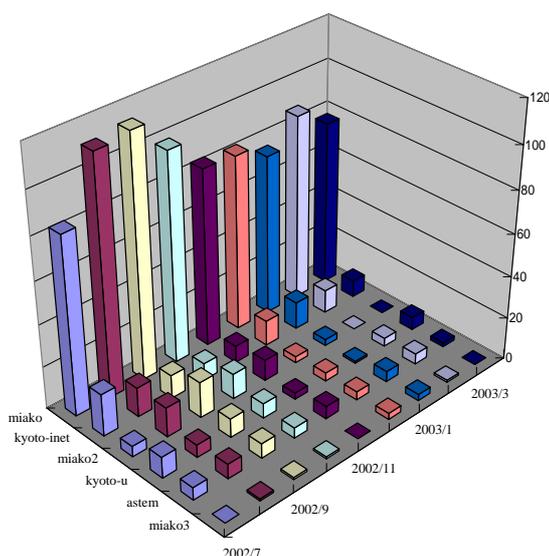
#### E. 付加価値

先述のとおり VoIP 市場は非常に巨大なマーケットになりつつあり、既存電話に比べ料金が安いだけでは競争力が低い。また現状で実用的に問題ない音質、遅延速度が達成できているため、VoIP 技術だけで差別化することは難しい。そこで何らかの付加価値をつける必要がある。

当該システムでは音声合成/認識システムを用いたユニバーサルデザイン、非同期システムを用いた音声メッセージング等、ユーザビリティ/アクセシビリティの高いニュースリーダーシステムなど、既存の VoIP システムにはない斬新な設計により付加価値を高めている

上記のように、当該システムでは VoIP 市場ニーズにおける、標準化の策定、安全性の追及、付加価値による競争力の強化を目指しつつ、P2P 型通信を用いることで既存技術に比べ低コストでシステムを構築することが可能である。

## [開発、実証中の miako-net 利用状況]



## 既存技術と比較した技術的優位性の評価

本方式 (NOTASHIP 方式) は、PDA 端末から PDA 端末、および携帯電話から PDA 端末を呼び出すのに、固定グローバル IP アドレスだけで、相手呼び出せることを確認・実証した。

また、PDA から携帯電話は、固定グローバル IP アドレスを割り当てられた Gateway を、

固定グローバル IP アドレスで呼び出し、ダイヤルトーンを送ることで携帯電話を呼び出すことができることを確認・実証した。

以上の結果より、

既存に利用されている SIP 方式と比較すると、呼設定 / 呼制御のためのゲートキーパは不要であることを実証した。

## 市場ニーズにおける技術コストの評価

シンプルでゲートキーパというセンターサーバ設備を必要としないため、低廉な通信料金で実現が可能である。

利用頻度の考察、

- (1) 利用頻度の高いところは、駅、商店街、交流会館などの人の往来が多くて、椅子に座って利用できるエリア
- (2) 旅館、喫茶店での利用が、意外と少ない
- (3) 屋外で椅子のないところに関しては、利用エリアの明確な表示がないこともあり、利用頻度が少ない

といえる。

また、利用できるエリアであるということを、いかに把握してもらうかが、重要であることもわかる。経済性の評価としては、

人の往来が多く、無料で利用できるエリアに入れ、椅子などに座ってゆっくりとできるところが、最も高いところと言える。

月次利用時間(秒)

			1月	2月	3月
屋内	椅子あり	駅 a	557,934	363,360	645,156
		駅 b	72,342	49,535	78,363
		駅 c	11,934	9,984	9,842
		商店街 a	169,732	459,334	533,768
		商店街 b	77,473	77,785	116,631
		大学	6,776	427	18,259
		旅館 a	1,266	493	5,331
		旅館 b	3,761	5,993	3,406
		喫茶店 a	168	2,011	2,474
		喫茶店 b	3,949	6,978	1,293
		交流会館 a	86,119	53,602	28,665
		交流会館 b	29,453	1,274	1,360
		福祉施設	32,680	16,216	81,814
		屋外	椅子あり	商業施設 a	552
商業施設 b	1,236			537	4,248
椅子なし	蛸薬師通		1,443	716	2,040
	室町通		2,246	105	910
	鴨川 a		3,918	2,422	845
	鴨川 b		259	242	202
	河原町通 a		332	976	664
	河原町通 b		820	4	0
	錦通 a		46	413	184
	錦通 b		0	27	1,280

## 2.音声認識/合成の評価実験

### 2.1 音声認識率の評価実験

#### [実験概要]

電話帳検索機能に実装された音声認識システムを用いて、認識率を計測する。

#### [計測方法]

50人分の氏名データ(被験者がわかりやすいように歴史上の人物名を使用)を用意し、その中から20分の氏名をランダムに抽出して発話。正常に認識が行われたかどうか確認する(認識が行われた場合結果ガイダンスされる)。

発話内容と同データが認識されたものを正常値とし、誤認識、未認識(10s以上たっても結果ガイダンスされない)ものを除き、1回の発話での測定(言い直しはしない)を行う。

30名の被験者を用いて計測を行い、場所、騒音状況などに関しては特に指定はしない。(ただし音声ガイダンスのエコーバックを避けるため、必ずヘッドフォンを着用)

#### [目標値]

85%以上の認識率

#### [実施内容]

##### 計測結果

被験者数(A)	30名
1被験者あたりの試験データ数(B)	20件
全試験データ数(A*B)	600件
正常認識データ数(C)	521件
認識異常データ数(D)	79件
認識率平均 : (C)/全データ数)*100	86.8%

#### [実施結果]

認識率は約87%であり、目標値をクリアしている。

通常、音声認識システムに関する技術的理解がある場合、意識的に音声認識しやすい発話を行う場合が多い。当該実験においても実験の趣旨はあらかじめ説明しているため、意識的に発話したため認識率が上がったことが考えられる。

しかし、当該実験は特に静穏な環境で行っておらず、むしろ音声認識には不利な日常環境でおこなったことを考慮すると、PDAという非力な環境で85%以上の認識率を達成できたことは、当該システムの音声認識機能が十分実用的であるといえる。

## 2.2 ユーザビリティの評価実験

### [実験概要]

音声認識における情報検索の評価を行うため、電話帳検索処理における応答時間を計測する。また、音声合成/認識システムを用いたシステムのユーザビリティに関する評価をインスペクション法を用いて行う。

### [計測方法]

2.1 と同じデータ、被験者を用いて音声認識が正常に行われた場合のみ、発話から認識結果ガイダンスがされるまでの時間を計測する。

計測は音声認識モジュール内に認識開始から認識終了(誤認識含む)までの時間を計測するプログラムスレッドを用意し、認識結果ガイダンス時に計測結果を画面表示した。

またインスペクション法によるユーザビリティの評価は 1.2 と同様の方法で行う。

### [目標値]

2s 以内の応答、10s 以内の検索

### [実施内容]

音声認識応答時間

計測結果



図.実験風景

被験者数(A)	30 名
1 被験者あたりの試験データ数(B)	20 件
全試験データ数(A*B)	600 件
正常認識データ数(C)	521 件
正常認識時の検索時間合計(D)	961.5s
検索時間平均 : (C)/(D)	1.845s

### ユーザビリティ評価

以下、ヒューリスティック評価項目に基づく評価内容を記載。

1.システム状態の視認性を高める	従来の画面表示に加え、音声合成による全ガイダンスを付加することにより、聴覚による視認性を高めたことは評価できるが、逆に視覚
------------------	---

	だけに特化した機能(フォントサイズの変更等)の実装が足りなかった部分もある。
2.実環境に合ったシステムを構築する	合成音質に関して好みがわかれる部分があるが、ある程度の調節が可能なことで、大多数のユーザをカバーできる。
3.ユーザにコントロールの主導権と自由度を与える	全体的に見て操作性が統一されていることが=自由度の低下につながる場合がある。しかし、それは視覚によるオペレーションを行う場合であり、当該システムにおいてはユーザビリティ向上のための自由度の低下は容認できる範囲である。
4.一貫性と標準化を保持する	システム全体としての操作性、応答性に関して一貫性が保たれており、システムの修得性を上げている。
5.エラーの発生を事前に防止する	通信に関して予期しないエラー(基地局のダウン)等が発生した場合、(すなわち自システムではなく、外部要因におけるエラーの場合)が発生した場合の防止手段がない。
6.記憶しなくても、見ればわかるようなデザインを行う	操作性の統一、機能の一般化により修得性は非常に高い。画面表示に関しても、現在の視覚重視なデザイン手法から見れば貧相であるが、逆にそのことが見ればわかるデザインにつながっている。
7.柔軟性と効率性を持たせる	1画面に対し1機能を割り当てたことにより、機能の拡張が容易であり効率的である。
8.最小限で美しいデザインを施す	画面表示デザインという意味でなく、ユーザビリティデザインという意味で考えれば、機能デザインとしての統合が取れている。
9.ユーザによるエラー認識、診断、回復をサポートする	5.と同等で外部要因におけるエラー発生時の診断と回復は難しい。
10.ヘルプとマニュアルを用意する	アプリケーションレベルで音声ガイダンスが用意されており、そのままヘルプとして利用できる。

#### [実施結果]

平均応答/検索時間は1.84秒であり、目標値をクリアしている。

ユーザビリティに関しては、音声合成による全表示項目のガイダンスという点から、必然的に1画面を1機能で完結させるというデザインになったが(複数ウインドウを開いたりメニューに様々な機能を割り付けない)、そのことがインターフェースの単純化と操作性の向上に影響を及ぼしたといえる。

全体的に見て、音声合成/認識インターフェースを用いたアプリケーション開発という点において、非常に高いユーザビリティを達成できたと考える。

## 2.3 アクセシビリティの評価実験

### [実験概要]

音声認識、合成技術を用いたシステム全体の操作性、到達性に関して、利用者アンケートによる評価実験、及び、入力等の操作に関して2つ以上のインターフェースが提供されているかどうかインスペクション法により評価する。

### [計測方法]

30人の被験者を用いてシステム全体の操作性、到達性に関しての評価調査を行う。調査は実験趣旨及びシステムの概要を説明した上で、使用していただき(希望者には1週間PDAを貸し出し)用意したアンケート項目に回答していただいた。被験者に関しては晴眼者の他に視覚障害者の方にも参加していただく(この場合アンケートには口頭で回答していただき、こちらで記入した)。また、入力、検索などの複雑な操作に関しては、アクセシビリティ向上の為、複数以上の操作方法を用意することで、特定環境下でのみの操作に依存しないことが重要である。その点に着目してインスペクション法によるアクセシビリティ評価を行った。

### [目標値]

全オペレーションに音声合成/認識インターフェースが用いられ、かつ入力等の操作に関して2つ以上のインターフェースが提供されていること

### [実施結果]

#### ユーザ評価

操作性、修得性に関しては概ね良好な結果が得られた。特に晴眼者と視覚障害者の間で評価的に大きな開きが無かったことは、アクセシビリティの向上という点において重要である。

中には今すぐ製品化してほしいという意見もあり、潜在的な需要が大きいことが感じとれた。

同等のインターフェースを用いた他機能の実装にも、参考となる意見をいただいたため、今後の課題としたい。

#### インスペクション法によるアクセシビリティ評価

以下の項目に関して評価を行った。

機能	音声インターフェース	複数オペレーション
VoIPシステム	全画面、全機能に関して音声合成によるガイダンスがされている。 また電話帳検索において音声	電話帳入力(文字入力)に関して、PDAハードウェアキーによる操作のほかに、PCで作成したデータをシンクロすることも

	合成インターフェースが用いられており、全オペレーションを音声のみで行うことが可能である。	可能であり、 <ul style="list-style-type: none"> <li>・PDA からの入力作業</li> <li>・PC からの入力作業</li> </ul> の2通りのオペレーションが提供されている。電話帳更新・削除に関しても同等である。 また、電話帳検索に関しては、 <ul style="list-style-type: none"> <li>・音声認識</li> <li>・50音検索</li> <li>・着発信履歴</li> </ul> の3通りのオペレーションが提供されている。
非同期システム	全画面、全機能に関して音声合成によるガイダンスがされている。 また電話帳検索において音声合成インターフェースが用いられており、全オペレーションを音声のみで行うことが可能である。	電話帳入力(文字入力)に関して、PDA ハードウェアキーによる操作のほかに、PC で作成したデータをシンクロすることも可能であり、 <ul style="list-style-type: none"> <li>・PDA からの入力作業</li> <li>・PC からの入力作業</li> </ul> の2通りのオペレーションが提供されている。電話帳更新・削除に関しても同等である。 また、電話帳検索に関しては、 <ul style="list-style-type: none"> <li>・音声認識</li> <li>・50音検索</li> <li>・着発信履歴</li> </ul> の3通りのオペレーションが提供されている。
ニュースリーダーシステム	全画面、全機能に関して音声合成によるガイダンスがされており、全オペレーションを音声のみで行うことが可能である。	入力等操作が無いため評価外

また、全画面全機能通して、ハードウェアキー操作異常時のスタイラスペン操作による機能終了操作も提供されていることから、目標値を達成できたと考える。

### 3.非同期システム

#### 3.1 アベイラビリティの評価実験

##### [実験概要]

ネットワークシステムにおける信頼性はダウンタイム(システム故障停止時間)を元にしたアベイラビリティによって検証されるため、1週間(168h)の連続稼働実験を行い、通常利用及び1日約2hの高負荷(10人以上同時利用)での故障及び修復時間を測定、また高負荷時におけるメモリ、ヒープ使用率、ネットワークトラフィックを計測し、ソフトウェア不具合による予測可能な故障時間を除く、予測可能な故障時間を付加した上で、システムの稼働率を検証した。

##### [計測方法]

10人分のHTTPリクエストを2h送り続けダウンタイムを計測  
168h(1週間)の計測を行い、アベイラビリティを算出

##### [算式]

$MTBF = (\text{稼働時間} - \text{停止回数} * \text{停止時間}) / \text{停止回数}$

$MTTR = (\text{停止回数} * \text{停止時間}) / \text{停止回数}$

$\text{アベイラビリティ} = MTBF / (MTBF + MTTR)$

##### [目標値]

99.9%以上の稼働率

##### [実施内容]

非同期システムにおける各サーバ機器に対して高負荷をかけるためのダミークライアントプログラムを作成。同時10人利用を想定し、ダミープログラムより10人分のリクエストを順に2h送り続け、その際のシステム情報を取得。同実験を1週間連続で行い、実停止時間とシステム情報により予測可能な故障時間を付加した結果を計測した。

なお、計測に関しては同一サーバを各2台用意した並列環境で行っている。

##### [実施結果]

計測内容

	稼働時間	停止回数	停止時間	修復時間	
Webサーバ1	168	2	0.5	1	
	MTBF				83.5
	MTTR				0.5
	アベイラビリティ				0.994048

Webサーバ2	稼働時間	停止回数	停止時間	修復時間	
	168	1	0.75	0.75	
	MTBF				167.25
	MTTR				0.75
	アベイラビリティ				0.995536

APPサーバ1	稼働時間	停止回数	停止時間	修復時間	
	168	2	0.5	1	
	MTBF				83.5
	MTTR				0.5
	アベイラビリティ				0.994048

APPサーバ2	稼働時間	停止回数	停止時間	修復時間	
	168	1	0.75	0.75	
	MTBF				167.25
	MTTR				0.75
	アベイラビリティ				0.995536

TELサーバ1	稼働時間	停止回数	停止時間	修復時間	
	168	6	0.25	1.5	
	MTBF				27.75
	MTTR				0.25
	アベイラビリティ				0.991071

TELサーバ2	稼働時間	停止回数	停止時間	修復時間	
	168	4	0.5	2	
	MTBF				41.5
	MTTR				0.5
	アベイラビリティ				0.988095

停止回数、停止時間は予測値含む

#### 算式

システム全体のアベイラビリティ

同一サーバ間 = 並列接続 -> 積

各サーバ間 = 直列接続 -> 和

#### 計算結果

Webサーバの並列構成故障率

$$(1 - 0.994047619) * (1 - 0.995535714) = 0.000026573129252$$

W1

W2

WW

APP サーバの並列構成故障率

$$(1 - 0.994047619) * (1 - 0.995535714) = 0.000026573129252$$

A1                      A2                      AA

TEL サーバの並列構成故障率

$$(1 - 0.991071429) * (1 - 0.988095238) = 0.000106292517007$$

T1                      T2                      TT

アベイラビリティ =

$$100 - (((WW + AA + TT) - (W1 * W2 * A1 * A2 * T1 * T2)) * 100)$$

= 約 99.98%

上記により、システムの平均稼働率は机上値で 99.98%であり、目標値を達成したことを実証した。

### 3.2 ユーザビリティの評価実験

#### [実験概要]

非同期システムのユーザビリティに関して、インスペクション法による評価及び、ユーザ登録等 Web ブラウザを使用する箇所に関して W3C で勧告される Web コンテンツにおけるユーザビリティを満たしているか評価する。

#### [実験内容]

ヒューリスティック評価項目に基づく評価内容を記載。

1.システム状態の視認性を高める	PDA クライアントに関しては音声ガイダンスが行われているため問題なし。ユーザ登録等 Web クライアントに関しては、スクリーンリーダでの音声読み上げに対応している為、オペレーショナルな問題は無い。
2.実環境に合ったシステムを構築する	PDA クライアントの利用環境を想定し、音声メッセージ登録時間に制限を設ける等、システムリソースや通信トラフィックの軽減につながる工夫がなされている。
3.ユーザにコントロールの主導権と自由度を与える	GUI ベースの入力アクションに対する応答を行うアプリケーションではない為、機能的にコントロールの自由度は少ない。 しかし、逆にそれ故に簡素な操作体系を実現している為、コントロールの不自由度=ユーザビリティの向上につながっている。

4.一貫性と標準化を保持する	操作性に関して一貫性が保たれている。
5.エラーの発生を事前に防止する	通信に関して予期しないエラー(基地局、非同期サーバのダウン)等が発生した場合、が発生した場合の防止手段がない。
6.記憶しなくても、見ればわかるようなデザインを行う	他機能と操作性が統一されている。また電話帳検索等はVoIPと同等のモジュールを使用しているため、統一性は高く、習得も容易である。
7.柔軟性と効率性を持たせる	音声メッセージの送受信にXMLベースの拡張フォーマットを用いることにより、同一の音声メッセージデータをPDA以外の様々なデバイスから利用する(PC、電話等)ことも可能であり、柔軟性は高い。
8.最小限で美しいデザインを施す	ユーザビリティデザインとしての最小限の機能的デザインがなされている。
9.ユーザによるエラー認識、診断、回復をサポートする	5.と同等で外部要因におけるエラー発生時の診断と回復は難しい。
10.ヘルプとマニュアルを用意する	アプリケーションレベルで音声ガイダンスが用意されており、そのままヘルプとして利用できる。

#### W3C Web コンテンツにおけるユーザビリティデザイン

No	項目	対応
1	すべての非テキスト要素に対してテキスト等価物を提供する	
	サーバー側イメージ・マップのそれぞれのアクティブ領域について、多様なテキスト・リンクを提供する	
	ユーザー・エージェントがマルチメディア表現の視覚的トラックのテキスト等価物を自動的に音読できるようになるまで、視覚的トラックの重要な情報の聴覚的説明を提供する	
	あらゆるタイムベースのマルチメディア表現(例、動画またはアニメーション)について、等価代替物(例、視覚的トラックのキャプションまたは聴覚的説明)を表現と同期させる	-
	ユーザー・エージェントがクライアント側イメージ・マップ・リンクに対するテキスト等価物をレンダリングできるようになるまで、クライアント側イメージ・マップでそれぞれのアクティブな領域に対して多様なテキスト・リンクを提供する	-
	色が表す全部の情報が、色なしでも(例、コンテキストまたはマークアップから)入手可能なことを保証する	

前景色と背景色の組合せが、色に関する障害を持つ人々が見たり、白黒画面で見た場合に十分なコントラストを提供することを保証する	
情報を表すのに適切なマークアップ言語が存在する場合、イメージではなく、マークアップを使用する	
発行された正式な文法に従った文書を作成する	
レイアウトや表現を制御するには、スタイルシートを使用する	
マークアップ言語の属性値とスタイルシートのプロパティ値には、絶対単位ではなく、相対単位を使用する	
文書構造を表すにはヘッダー要素を仕様に従って使用する	
リストおよびリスト項目は適切にマークアップする	-
インデントなどの効果をフォーマットするために引用マークアップを使用しない	-
文書のテキストおよび任意のテキスト等価物(例、キャプション)の自然言語の変更を明確に識別する	
文書の中で省略語または頭字語が最初に現れたとき、そのもとのスペルを指定する	-
文書の主となる自然言語を識別する	
データ・テーブルの場合、行と列のヘッダーを識別する	
行または列ヘッダーに複数の論理レベルがあるデータ・テーブルの場合、マークアップを使って、データ・セルとヘッダー・セルを関連付けする	
直線化したとき意味を持つ場合を除き、テーブルをレイアウトのために使用しない	
テーブルをレイアウトのために使用する場合、構造マークアップを視覚的フォーマットの目的で使用しない	
テーブルの要約を提供する	-
ヘッダー・ラベルの省略語を提供する	-
文書は、スタイルシートなしでも読めるような構成にする	
動的コンテンツが変更される場合、それに対する等価物も更新されることを保証する	
スクリプト、アプレット、その他のプログラマティック・オブジェクトがオフになっていたたり、サポートされていないときでも、ページが使用可能であることを保証する	
動的コンテンツがアクセス可能であるか、代替の表現またはページを提供することを保証する	
ユーザー・エージェントを使ってユーザーがフリッカーを制御できるようになる	-

	まで、画面のフリッカーは避ける	
	ユーザー・エージェントを使ってユーザーが点滅を制御できるようになるまで、コンテンツの点滅は避ける	-
	ユーザー・エージェントを使ってユーザーが移動するコンテンツを止めることができるようになるまで、ページ内の移動は避ける	-
	ユーザー・エージェントがリフレッシュを停止する機能を提供するまで、ページの定期的な自動リフレッシュは行わない	
	ユーザー・エージェントが自動リダイレクトを停止する機能を提供するまで、マークアップを使って、ページを自動的にリダイレクトしない	
	スクリプトやアプレットなどのプログラマチックな要素は、直接アクセス可能にするか、支援技術に対する互換性を保証する	
	領域が利用可能な幾何学形状によって定義できない場合を除き、サーバー側のイメージ・マップではなく、クライアント側のイメージ・マップを提供独自のインターフェースを持つ要素は、デバイスに依存しない方法で動作できるように保証する	-
	スクリプトの場合、デバイスに依存するイベント・ハンドラではなく、論理イベント・ハンドラを指定する	
	リンク、フォーム・コントロール、オブジェクトを通じて論理的なタブ順序を作成する	
	重要なリンク、フォーム・コントロール、フォーム・コントロールのグループへのキーボード・ショートカットを提供する	
	ユーザー・エージェントを使って、ユーザーが新しく作成されたウィンドウをオフにできるようになるまで、ポップアップまたはその他のウィンドウを表示しない	
	ユーザー・エージェントがラベルとフォーム・コントロール間の明示的な関連付けをサポートするまで、ラベルが暗示的に関連付けられたすべてのフォームについて、ラベルが適切に位置決めされることを保証する	
	ユーザー・エージェントがテキストを正しく並べてレンダリングできるようになるまで、テキストを並行にレイアウトするすべてのテーブルについて、直線的なテキスト代替物を提供する	
	ユーザー・エージェントが空のコントロールを正しく取り扱うことができるようになるまで、編集ボックスとテキスト領域には、デフォルトのプレースホルダー文字をいれる	
	ユーザー・エージェントが隣接するリンクを明確にレンダリングできるようになるまで、隣接するリンクの間に非リンクの印刷可能文字をいれる	
	W3C 技術が利用可能で、かつ作業にとって適切な場合、それを使用し、最新	

のバージョンがサポートされている場合、それを使用する	
廃止された W3C 技術は避ける	
ユーザーが自分の好みに従って文書を受け取れるように情報を提供する	-
最善の努力を払ってもアクセス可能なページが作成できない場合、W3C 技術を使用し、アクセス可能で、等価な情報を持ち、アクセス不可能なページと同じ頻度で更新される代替ページへのリンクを提供する	
フレームの識別とナビゲーションを容易にするために、各フレームに表題を付ける	-
フレームの目的とフレーム間の関係が、フレーム表題だけではわかりにくい場合、それらを説明する	-
大きな情報ブロックは、自然で適切な、より取り扱いやすいグループに分割する	
ラベルは明示的にコントロールと関連付ける	
各リンクのターゲットを明確に識別する	
ページやサイトに意味情報を追加するメタデータを提供する	
サイトの一般的なレイアウトに関する情報を提供する	
ナビゲーション機構は一貫性のある方法で使用する	
ナビゲーション機構を強調表示し、それに対するアクセスを与えるためのナビゲーション・バーを提供する	
関連するリンクをグループ化し、グループを識別し、ユーザー・エージェントがこれらを行うまでは、グループをバイパスするための方法を提供する	
検索機能を提供する場合、各種の技能レベルや設定に合った複数の種類の検索を可能にする	
見出し、パラグラフ、リストなどの最初に区別情報を配置する	
文書コレクションに関する情報を提供する	-
複数行の ASCII アートをスキップするための手段を提供する	-
サイトのコンテンツにとって最もわかりやすく、シンプルな言語を使用する	
グラフィックまたは聴覚的表現は、ページの理解に役立つ場合、テキストを補う	
ページ間で一貫性のある表現のスタイルを作成する	

問題なし 一部問題あり ×問題あり 該当なし

#### [実施結果]

上記結果より、ISO13407 で定義されるヒューマンインターフェースデザインに準拠し、また W3C で勧告される Web コンテンツにおけるユーザビリティデザインについても準拠することを確認した。

### 3.3 スケーラビリティの評価実験

3.1 と同一条件での稼働実験を行い、システムの応答時間を計測する。また高負荷時における分散前と分散後のパフォーマンスを計測し、負荷分散によるスケーラビリティの確保が容易であるか検証する。

#### [計測方法]

単一システムで 25 人分の HTTP リクエストを順に 2h 送り続け応答時間を計測

計 1 週間の計測を行う。

次に負荷リクエスト数を 50 人分に増やし、分散システム(各サーバ 2 台の並列動作)での応答時間を計測。

#### [目標値]

負荷分散時に 2s 以内の応答時間

#### [実施内容]

3.1 と同等のダミークライアントを用いて、HTTP Get による AVML 取得リクエスト(システムでは未読/既読の音声メッセージ一覧取得に使われるリクエスト)を 25 リクエスト(25 人分)発行(1s 間隔で発行)し、単一システム(各サーバ 1 台構成)での応答時間を計測し、平均応答時間が 2s 以内かを検証。

次に、リクエスト数を 50 に増やし、同時に各サーバを 2 台構成にし負荷分散を行いながら、同等のリクエストによる応答時間を計測。

単一システムと分散システムのパフォーマンスを比較し、高負荷時の負荷分散が効率的に行われているかどうかを検証した。

なお、上位ネットワークトラフィック(プロバイダ等)によって計測結果が左右されないよう、ネットワークは LAN 構成(各接続は 100Mbps)で行った。(したがって DNS は使用せず、直接 IP アドレスでサーバを指定)

#### [実施結果]

##### 計測内容

単一構成		分散構成			
リクエスト	平均応答時間	リクエスト	平均応答時間	リクエスト	平均応答時間
1	1.43	1	1.89	26	1.96

2	1.27	2	1.94	27	1.93
3	1.09	3	1.76	28	1.89
4	1.24	4	1.91	29	2.02
5	1.32	5	1.93	30	1.97
6	1.64	6	1.86	31	1.95
7	1.2	7	1.91	32	1.99
8	1.36	8	1.87	33	2.03
9	1.49	9	1.84	34	1.96
10	1.13	10	1.9	35	1.99
11	1.42	11	1.92	36	2.02
12	1.48	12	1.92	37	2.13
13	1.41	13	1.87	38	2.04
14	1.39	14	1.91	39	1.95
15	1.44	15	1.86	40	1.99
16	1.47	16	1.89	41	2.03
17	1.49	17	1.92	42	2.16
18	1.52	18	1.89	43	2.23
19	1.49	19	1.88	44	2.24
20	1.51	20	1.91	45	2.21
21	1.54	21	1.93	46	2.15
22	1.5	22	1.91	47	2.18
23	1.52	23	1.84	48	2.31
24	1.59	24	1.86	49	2.33
25	1.56	25	1.93	50	2.25
平均	1.42				1.9832

結果的には目標値はクリアしたものの課題の残る結果となった。予想値では単一構成の場合と同等の応答時間が期待されたが(同程度のパフォーマンスが出ないならクラスタリングする必要性が少ないため)、予想値を超えるオーバーヘッドが見られた。

原因としてメインロジックを行うアプリケーションサーバでは効率的にリソースの共有が行われていたが、ネットワーク I/O を行う Web サーバが問題と思われる。

Web サーバは各クライアントとの I/O 用にセッション情報を持つが、現在 Web サーバで使用している Servlet コンテナである Tomcat が、複数のコンテナ間でセッション情報を共有するセッションデータリポジトリを実装していないため、その部分に関してクラスタリングがなされていない。

したがって、各 Servlet 間で独自にセッションテーブルを持ち定期的にマージしているのだが、その部分のオーバーヘッドが計測結果に反映されたと思われる。

これは HTTPD である Apache と Servlet コンテナ Tomcat 間を CORBA/iiop によって実装することで、セッションデータリポジトリが可能になり、さらに効率のよい負荷分散が可能と思われる。

#### 4. 結言

各評価より、想定した目標値は達成できている。

特にユーザビリティに関しては、実証実験に参加していただいた方以外にも、盲学校でのデモンストラクション等で高評価をいただいております。今後、ユニバーサルデザインの理念に基づくソフトウェア開発の指針となる結果を得られたことは、大きな成果であるといえる。