

東京大学における iMac のディスクイメージ更新

品川 高廣

東京大学 情報基盤センター

shina@ecc.u-tokyo.ac.jp

Updating Disk Images of iMac in the University of Tokyo

Takahiro Shinagawa

Information Technology Center, The University of Tokyo

概要

東京大学情報基盤センターでは「教育用計算機システム (ECCS)」を運用しており、1,341 台の iMac が稼働している。これらの iMac 上では、Mac OS と Windows がデュアルブートで動作しており、CAD を含む様々なソフトウェアが快適に動作する環境を提供している。この環境では、OS やソフトウェアの更新などの際に各 iMac のディスクイメージを一斉に更新する必要があるが、100GB 前後のイメージ配信には非常に時間がかかる。本論文では、ECCS で 2016 年 3 月から導入された vThrii Seamless Provisioning について紹介する。このシステムでは、バックグラウンドで差分イメージ配信をすることで配信時間を短縮しつつ、更新が完了する前であっても新しい OS イメージで直ちに起動させることができる。また、万が一イメージ配信に失敗した場合は、もとのイメージに戻すこともできる。これにより、端末のイメージ管理の負担を大幅に削減することができるようになった。

1 はじめに

東京大学情報基盤センターでは、東京大学の学生及び教職員が教育・研究のために利用する「教育用計算機システム」(以下、ECCS = Educational Campuswide Computing System) を運用している。2016 年 3 月に導入された ECCS2016 では 1,341 台の iMac が稼働しており、駒場・本郷・柏の 3 つのキャンパスに分散配置されている。各 iMac 上では、Mac OS と Windows がローカルディスク (SSD) からデュアルブートするようになっており、CAD を始めとした各種アプリケーションが快適に動作する環境を提供している。

このような環境では、各 iMac のローカルディスクのイメージ管理が重要となる。まず、多数の iMac が物理的に分散配置されていることから、各 iMac のディスクイメージを一箇所で集中管理する仕組みが必須となる。また、各 iMac のローカルディスクの容量は 256GB であり、各 OS のイメージサイズは 100GB 程度となる。OS やソフトウェアを更新するためには、これらの分散した巨大なディスクイメージを定期的に一斉に更新する必要がある。しかし、ディスクイメージの更新にはイメージファイルネットワークで配信する必要があり、巨大なイメージの更新は非常に時間がかかる。

本論文では、東京大学の ECCS2016 で採用された vThrii Seamless Provisioning の機能について紹介する。vThrii Seamless Provisioning は、仮想化技術を活用することで、OS からは独立した形でディスクイメージをサーバからローカルディスクに配信することができる。また、まだ配信されていない領域へのアクセスがあった場合は、ネットワーク越しにサーバからデータを取得することができる。これにより、ディスクイメージ全体の配信が終了するのを待つことなく、配信途中であっても新しい OS を起動して動作させることが可能になる。

一方、未配信の部分は OS が稼働中であってもバックグラウンドで配信をおこない、最終的にはディスクイメージ全体を配信するようにする。これにより、配信終了時には全てのディスクアクセスがローカルディスクへのアクセスとなり、ネットワークの遅延に影響されない快適な環境を提供することが可能になる。更に、差分イメージのみを配信できるようにすることで、配信完了までにかかる時間を大幅に短縮することが出来る。

仮想化によるオーバーヘッドを最小限に抑えるために、vThrii のハイパーバイザはハードウェアの大部分を仮想化せず、ディスクイメージ配信に必要なストレージデバイスとネットワークデバ

イスを少しだけ仮想化する。これにより、ディスクイメージ配信終了後は、ハイパーバイザがほとんど動作しないようにして、端末の動作が遅くならないようにする。

なお、vThrii Seamless Provisioning は、東京大学情報基盤センターと株式会社イーゲルとの共同研究により研究開発した成果をベースとして製品化したものであり、その技術的詳細は国際会議 ASPLOS 2015 [1]において発表している。

以下では、まず 2 章で ECCS2016 の要件について述べ、3 章では従来方式の問題点について述べる。4 章で vThrii Seamless Provisioning の機能の詳細について述べ、5 章で運用実績を示し、6 章で本論文をまとめる。

2 ECCS2016 の要件定義

東京大学では、2016 年 3 月から ECCS2016 を導入するにあたり、以下のような要件を設けた。

- (1) Mac OS と Windows の両方が動作すること
- (2) SSD からローカルブートすること
- (3) 端末管理が容易であること

東京大学では 1 年生の必修の授業で Mac OS を標準環境として使用する一方、Windows 用の CAD ソフトウェアを使う授業などもあり、Mac OS と Windows の両方の環境が動作することが必須である。また、OS やアプリケーションの起動の速さや利用時の快適性を重視して、ネットワークの遅延による影響が出ることがないローカルブート方式を選択した。また、多数の端末が物理的に分散配置されていることから、ディスクイメージを一箇所で集中管理することも必須であった。

3 従来方式の問題点

従来 of 端末管理方式には、大きく分けるとシンクライアントとイメージ配信の 2 つの方式がある。シンクライアントには、ネットワークブートと画面転送型の 2 つの方式がある。ネットワークブートでは、サーバに置かれたディスクイメージをネットワーク経由で取得して起動するため、ディスクイメージを集中管理することができる。しかし、OS の稼働中のディスクアクセスが全てネットワーク経由となるため、ネットワークとサーバに負荷がかかり、端末の性能が低下することがある。

画面転送型には、サーバーベース方式、ブレード PC 方式、仮想 PC 方式などがあるが、いずれも

OS やアプリケーションをサーバ側で実行してその画面だけをクライアントに転送する方式である。この方式では、ディスクイメージがサーバ側に置かれることになるため、一箇所で集中管理することができる。しかし、画面がネットワーク経由で転送されるため、アプリケーションによっては応答速度が問題になる。また、アプリケーションとの互換性に問題が生じたり、サーバ側に高価な機器が必要であったり、仮想化のオーバーヘッドが生じたりするなどの問題が発生する可能性がある。

従来のイメージ配信でも、OS のディスクイメージをサーバからローカルディスクに配信することで、ディスクイメージの集中管理が可能になる。また、OS やアプリケーションからのディスクアクセスがローカルディスクに対しておこなわれ、実行中にネットワークアクセスが不要になるため、OS やアプリケーションの応答性は通常のマシンと同等の快適さを実現することができる。しかし、イメージ配信には非常に時間がかかり、配信中はマシンを使用することができない。また、万が一イメージ配信に失敗したときには、ローカルディスクから起動できなくなり、マシンが使えなくなる可能性がある。

4 vThrii Seamless Provisioning

vThrii Seamless Provisioning では、ディスクイメージの集中管理とローカルディスクアクセスの快適さを両立するために、ハイパーバイザが以下のような機能を提供している。

- (1) ネットワーク起動
- (2) 持続性キャッシュ
- (3) バックグラウンド導入
- (4) ディスク凍結
- (5) 差分配信
- (6) 巻き戻し

以下では、それぞれの機能について説明する。

4.1 ネットワーク起動

ネットワーク起動では、ディスクイメージをサーバからクライアントにネットワーク経由で転送して OS を起動する。これにより、ディスクイメージをサーバで集中管理できるほか、サーバで新しいイメージに変更したら、それをすぐにクライアントに適用することが可能であり、イメージ配信の完了を待たなくても良い。

4.2 持続性キャッシュ

持続性キャッシュでは、ネットワークから転送したディスクイメージをキャッシュとしてローカルディスクに保存する。これにより、一度転送したデータをサイドネットワークから転送しなくてもよくなる。例えば、2 回目の OS 起動はローカル起動と同等の速度で動作するようになる。また、よく使うデータから先にローカルディスクに格納されるため、あまり使わないデータの転送を後回しにすることが出来る。

4.3 バックグラウンド導入

バックグラウンド導入では、空き時間にハイパーバイザがサーバからディスクイメージを読み込んで、ローカルディスクに書き込む作業をおこなう。これにより、サーバのディスクイメージが徐々にローカルディスク全体に書き込まれてゆき、最終的にはネットワークへのアクセスが一切不要になる。

4.4 ディスク凍結

ディスク凍結では、ローカルディスクに対する書き込みを別領域に保存しておき、電源を OFF にするときその内容を破棄することによって、元のディスクイメージが書き換えられない状態（読み込みのみの状態）に保つ機能である。これにより、ウィルス感染や意図せぬ設定変更などがあっても、電源の再投入に依って元の状態に戻すことが出来る。

4.5 差分配信

差分配信は、ローカルディスクに既に前のディスクイメージが格納されているときに、新しいディスクイメージとの差分部分だけを送ることが出来るようにする機能である。これにより、ネットワーク経由で配信するデータの量を削減して、配信の完了に係る時間を短くすることが出来る。また、ディスク容量があれば、複数のバージョンを持たせることも可能になる。

4.6 巻き戻し

ロールバックでは、ディスクイメージを新しくする際に、元のディスクイメージを取っておいて、瞬時に戻すことが出来るようにする機能である。これにより、新しく作成したディスクイメージに

問題があることが発覚した場合でも、元のディスクイメージで即座に運用を継続することができるようになる。

5 運用実績

現時点では、およそ月 1 回のペースでイメージ更新をおこなっている。いずれも更新による性能低下を最小限に抑えるために、夜間に自動起動することで配信をおこなっている。差分更新ということもあり、いずれも夜間で十分に全てのイメージ配信が終了している。

なお、運用中に OS X のイメージに不備があることが発覚して、巻き戻し機能で1つ前のイメージに戻して対策をおこなったことが 1 回ある。また、特定の教室に緊急にソフトウェアをインストールするために、夕方にイメージ更新をしたことがあり、このときは 15 分程度でイメージの更新が完了した。

このように、現在のところイメージ更新によるユーザに対する影響は殆どない状態を維持しているほか、不測の事態に対しても対処する高い柔軟性と信頼性を有したシステムとなっている。

6 まとめ

本論文では、東京大学情報基盤センターが運用する ECCS2016 で導入された vThrii Seamless Provisioning の機能を紹介した。本システムでは、仮想化技術を活用して、バックグラウンドで差分イメージ配信をおこなうことにより、(1)配信完了を待つことなく新しい OS イメージで起動することが可能になり、(2)イメージ配信完了までの時間を短縮し、(3)最終的には仮想化やネットワークのオーバーヘッドを限りなく減らして、ローカル起動と同等の性能を実現することが出来るようになった。これにより、端末のイメージ管理の負担を大幅に削減することが可能になった。

参考文献

- [1] Yushi Omote, Takahiro Shinagawa, Kazuhiko Kato, Improving Agility and Elasticity in Bare-metal Clouds, In Proceedings of the 20th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2015), pp. 145-159, Mar 2015.