

Windows Server 2008 R2: Dynamic Memory

Андрей Бешков

Эксперт по ИТ-инфраструктуре, Майкрософт Россия

<http://blogs.technet.com/abeshkov>

<http://twitter.com/abeshkov>

abeshkov@microsoft.com

Пользователи, память и виртуализация

Мало кто может правильно выделить ресурсы для VM

Сколько действительно памяти требуется для IIS под конкретный сайт/проект?

- Принт-серверу?
- Файловому серверу?
- BranchCache?
- DirectAccess?

Как изменится производительность VM, если уменьшить память? Увеличить в двое? Втрое? Что изменится для других VM на том же хосте?

Есть желание проводить расчеты?

“Любая новая VM получает 1GB
[не зависимо от роли ее OS]. Память добавляется после анализа производительности”

“Всем VM по 4GB
[не волнует, что дальше будет с этой «лишней» памятью, проводить анализ накладно] и все довольны – кроме бизнеса, который за это платит”

“Устанавливаются минимальные или рекомендуемые параметры для ОС и приложений. Производитель всегда прав – есть ли смысл в тонкой настройке и тестировании?”

Виртуализация и память

- Память
 - Ключевой фактор плотности VM на оборудовании
 - Часто – самый дорогой компонент в системе
- Требования бизнеса
 - Повышение плотности VM при минимальном влиянии на производительность
 - Поддержка соответствующей производительности
 - Стабильное в продуктивном использовании решение

Технологии управления памятью VM

Memory Overcommit

- Overcommit (англ.) – невыполнимые обязательства
- Термин – маркетинговый, а не технический
- Означает «выделить больше памяти, чем существует физически». Технически может означать любую технологию:
 - Page Sharing
 - Second Level Paging
 - Dynamic Memory Balancing (a.k.a. ballooning)

Dynamic Memory HE Overcommit

- В реальности – никто не хочет использовать overcommit
 - Вы не можете «перебрать» ваши ресурсы – за все требуется «расплата»
 - VMware не рекомендует использовать ее технологии overcommit в продуктовой среде
- Dynamic Memory утилизирует память так же, как ядро ОС – CPU для многозадачности

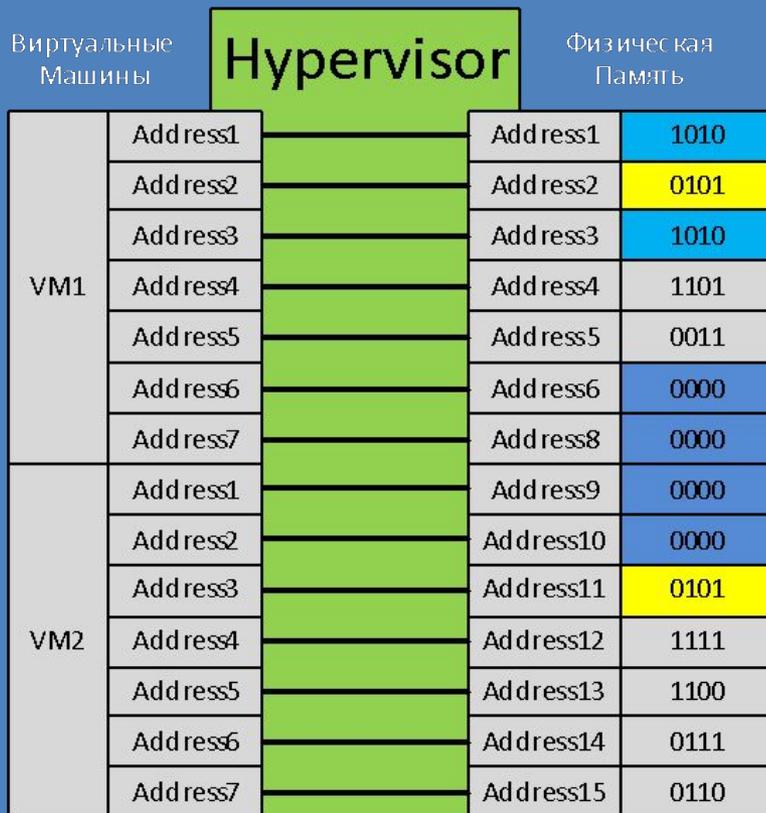
Page Sharing

Page Sharing

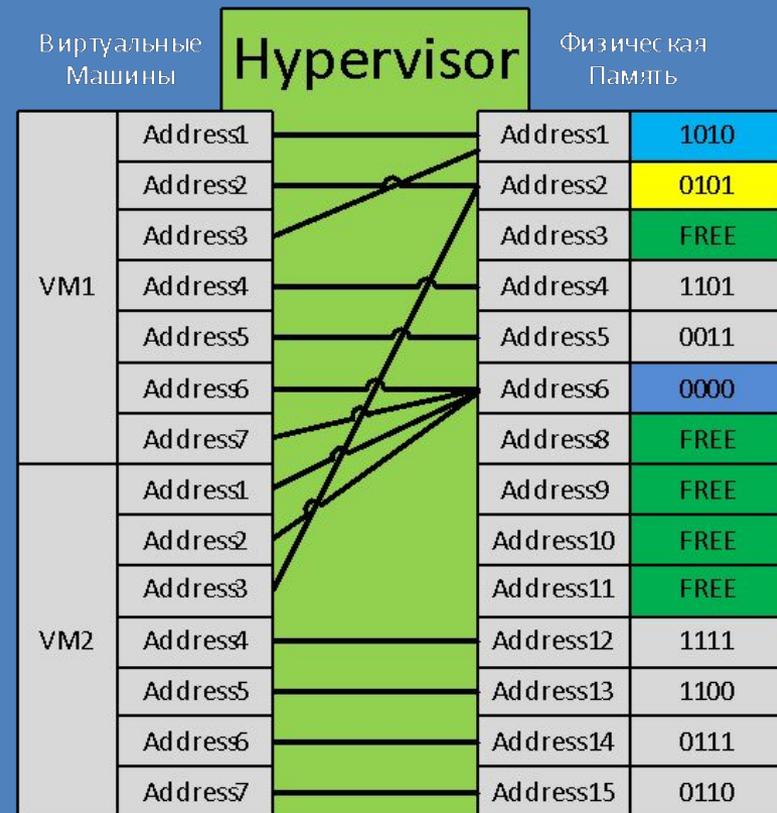
- Как это работает:
 1. Вычисляются контрольные суммы всех страниц памяти и заносятся в таблицу
 2. По всей таблице ищутся совпадающие значения
 3. Производится побитовое сравнение страниц с одинаковыми контрольными суммами
 4. Идентичные страницы хранятся в одном экземпляре

Page Sharing

No Page Sharing



Transparent Page Sharing



Эффективность Page Sharing

- О чем молчит маркетинг VMware?
 - Процесс расчета и сравнения контрольных сум долог и затратен (может занять несколько часов)
 - Page Sharing – не динамическая технология
 - Лучше всего «сжимаются» свободные (нулевые) страницы памяти

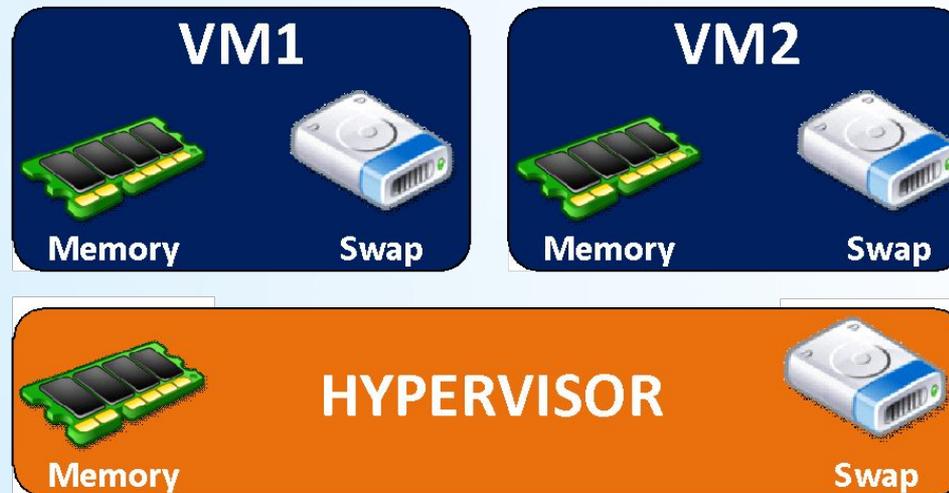
Эффективность Page Sharing

- Если не используются большие страницы памяти (2Мб)
 - 4Кб-страницы – это вчерашний день
 - Низкая производительность
- Если ОС не утилизирует память полностью (SuperFetch)
- ASLR – технология размещения исполняемого кода по случайным адресам – снижает эффективность Page Sharing

Second Level Paging

Second Level Paging

- Свопинг страниц памяти на жесткий диск на уровне гипервизора
- Гипервизор не знает, что происходит внутри гостевых ОС



Second Level Paging

Чем это плохо?

- Выгрузка в файл подкачки области памяти ядра гостевых ОС
- Двойная подкачка
- Падение производительности



Second Level Paging (1/3)

Проблема 1: Свопинг ресурсов ядра гостевой ОС

- *“the guest operating system will never page out its kernel pages since those pages are critical to ensure guest kernel performance. The hypervisor, however, cannot identify those guest kernel pages, so it may swap them out. In addition, the guest operating system reclaims the clean buffer pages by dropping them. Again, since the hypervisor cannot identify the clean guest buffer pages, it will unnecessarily swap them out to the hypervisor swap device in order to reclaim the mapped host physical memory.”*

- *–Understanding Memory Resource Management in VMware ESX Server p. 9-10;
<http://www.vmware.com/resources/techresources/10062>*

Second Level Paging (2/3)

Проблема 2: двойной пейджинг

- *“Assuming the hypervisor swaps out a guest physical page, it is possible that the guest operating system pages out the same physical page, if the guest is also under memory pressure. **This causes the page to be swapped in from the hypervisor swap device and immediately to be paged out to the virtual machine’s virtual swap device. Note that it is impossible to find an algorithm to handle all these pathological cases properly. ESX attempts to mitigate the impact of interacting with guest operating system memory management by randomly selecting the swapped guest physical pages.**”*
- *–Understanding Memory Resource Management in VMware ESX Server p. 9-10;
<http://www.vmware.com/resources/techresources/10062>*

Second Level Paging (3/3)

- Проблема 3: производительность памяти против диска
 - Доступ к памяти:
 - DDR3-1600 = 5 ns; DDR3-1333 = 6 ns
 - DDR3-1066 = 7.5 ns; DDR3-800 = 10 ns
 - Поиск данных на диске (disk seek): ~8 milliseconds
 - Формула сравнения DDR3-800 и диска: $.008 / .0000000010$
 - DDR3-1600 в 1,600,000 раз быстрее диска
 - DDR3-1333 в 1,333,333 раз быстрее диска
 - DDR3-1066 в 1,066,666 раз быстрее диска
 - DDR3-800 в 800,000 раз быстрее диска
 - Dual layers of paging
 - Чрезмерное употребление снижает вашу производительность на порядки...

Падение производительности

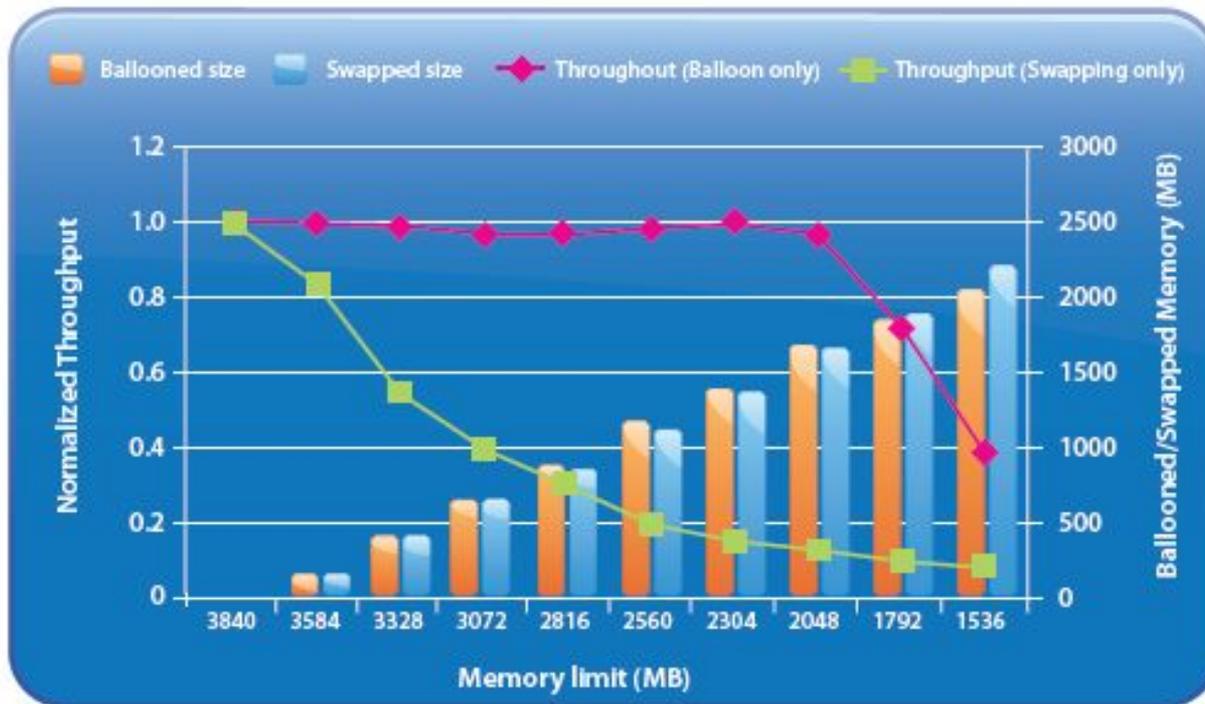
- -Understanding Memory Resource Management in VMware ESX Server p. 16;

<http://www.vmware.com/resources/techresources/10062>

5.3.2 Oracle/Swingbench

Figure 11 presents the throughput of an Oracle database tested by the Swingbench workload with different memory limits when using ballooning vs. swapping. The throughput is normalized to the case where virtual machine memory is not reclaimed.

Figure 11: Performance of Swingbench when using ballooning vs. swapping



Hyper-V R2 SP 1 Dynamic Memory

Цели Dynamic Memory

- Высокий уровень плотности VM с минимальным влиянием на производительность системы в целом (как хоста, так и гостевых ОС)
- Одинаково хорошо работать с различными типами нагрузок VM – серверами или десктопами, с микшированной нагрузкой
- Добавлять минимальную нагрузку на систему, особенно – память
- Проходить тест – «вроде, это выглядит правильно»

Запустили виртуалки

- Финансы
- Продажи
- Разработчики

8 GB
6 GB
4 GB
2 GB



3 VM
запущено

T = 0 T = 15 T = 30

Настройки
памяти
виртуалок

Memory management

Specify a set amount of memory for this virtual machine, or let Hyper-V amount dynamically within the specified range.

Static

RAM: MB

Dynamic

Startup RAM: MB

Maximum RAM: MB

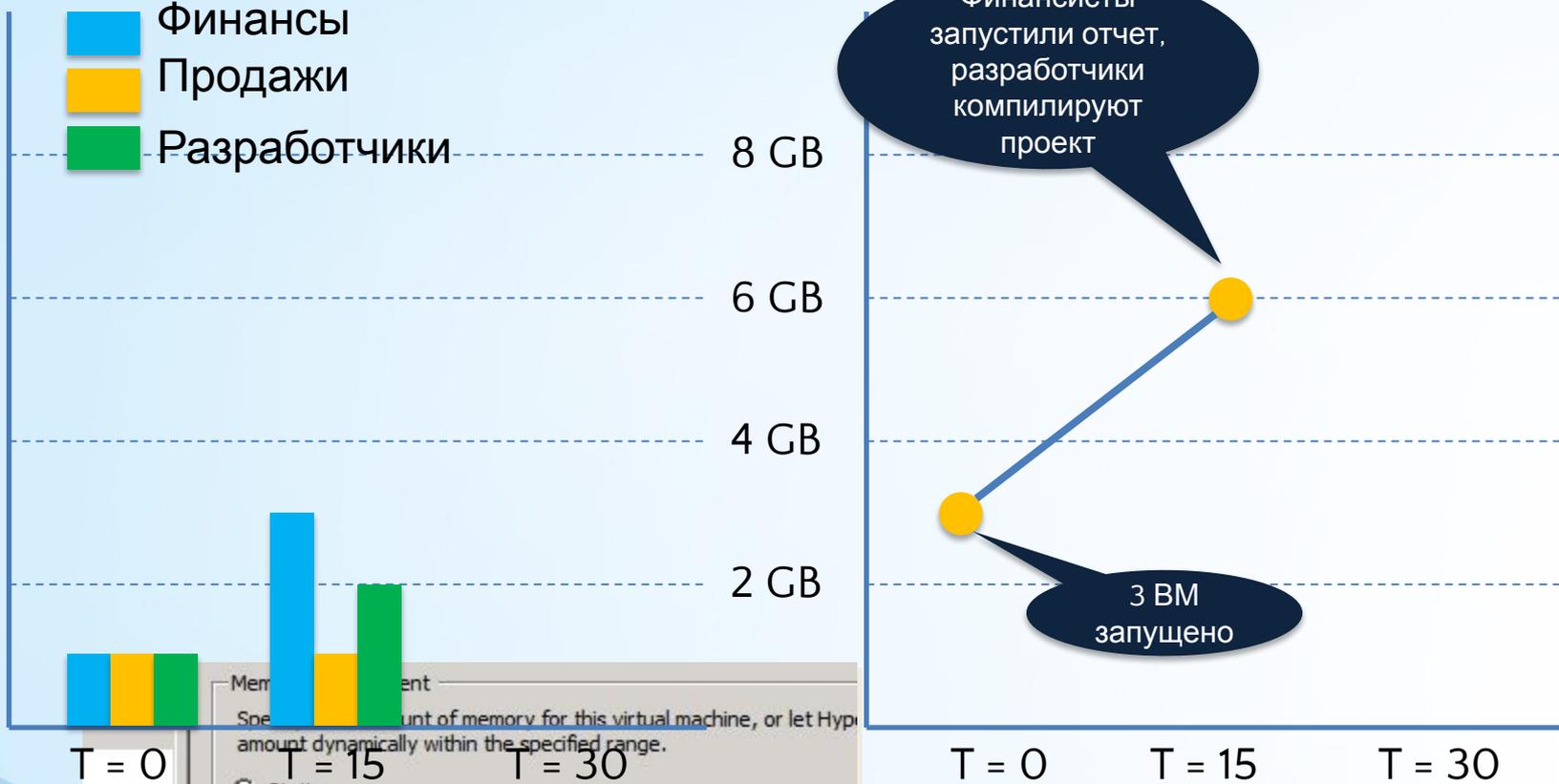
Specify the percentage of memory, based on the workload of the machine, that Hyper-V should try to reserve as a buffer.

5% 95%

всего памяти	8 GB
используется виртуалками	3 GB
фрагментация памяти	37.5%

Прошло 15 минут...

- Финансы
- Продажи
- Разработчики



Настройки памяти виртуалок

Memory settings for a virtual machine:

- Static:
- Dynamic:
- RAM: 512 MB
- Startup RAM: 1024 MB
- Maximum RAM: 4096 MB

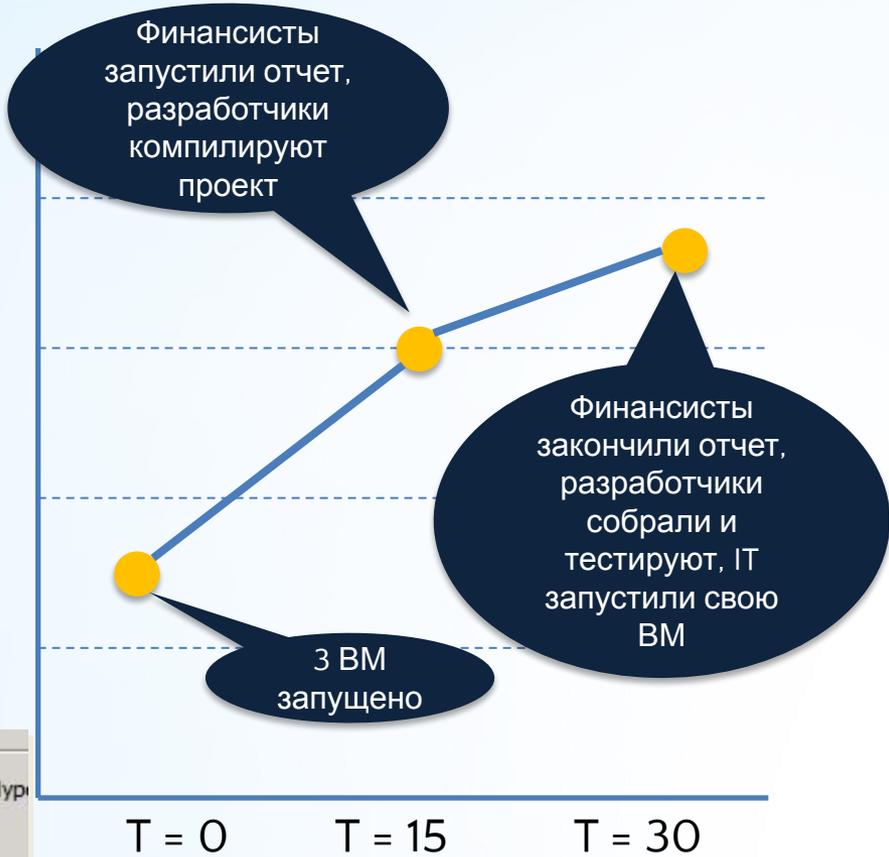
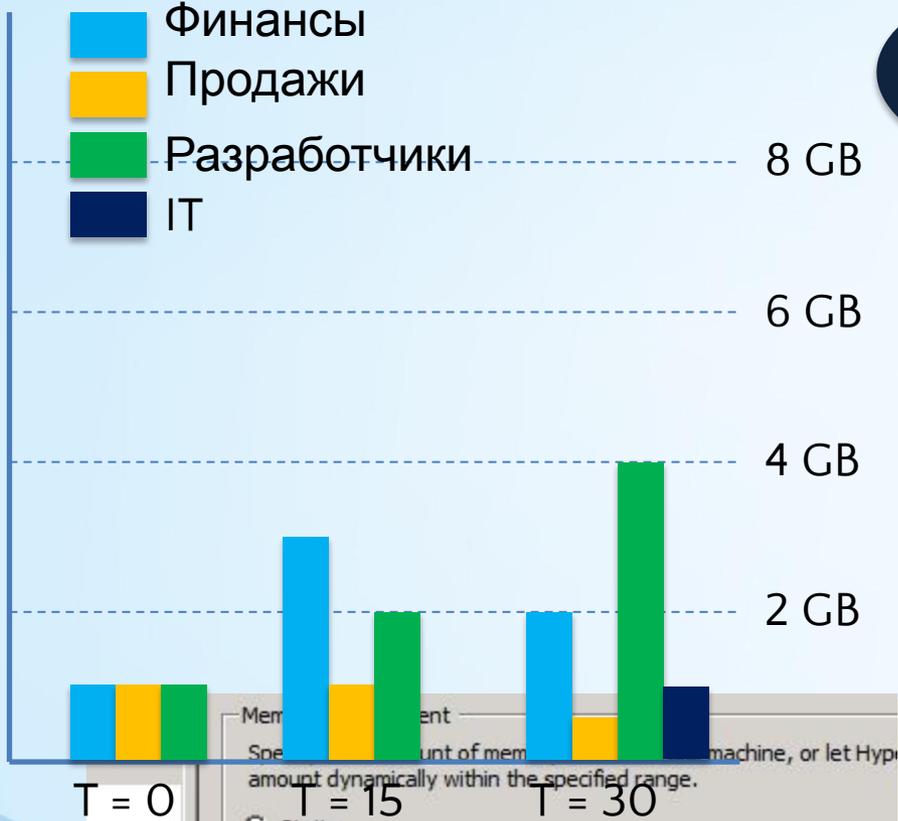
Specify the percentage of memory, based on the workload of the machine, that Hyper-V should try to reserve as a buffer.

5% 95%

Всего памяти	8 GB
Используется ВМ	6 GB
Утилизация памяти	75%

Прошло 15 минут...

- Финансы
- Продажи
- Разработчики
- IT



Настройки памяти VM

Memory settings for the VM:

- Static:
- Dynamic:
- RAM: 512 MB
- Startup RAM: 1024 MB
- Maximum RAM: 4096 MB

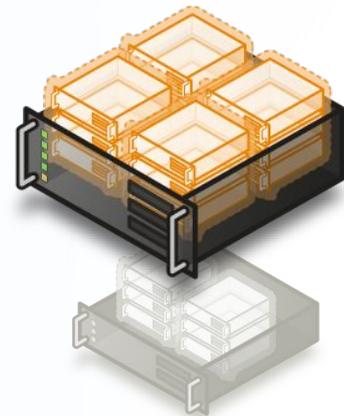
Specify the percentage of memory, based on the workload of the machine, that Hyper-V should try to reserve as a buffer.

5% 95%

Всего памяти	8 GB
Используется VM	7.5 GB
Утилизация памяти	94%

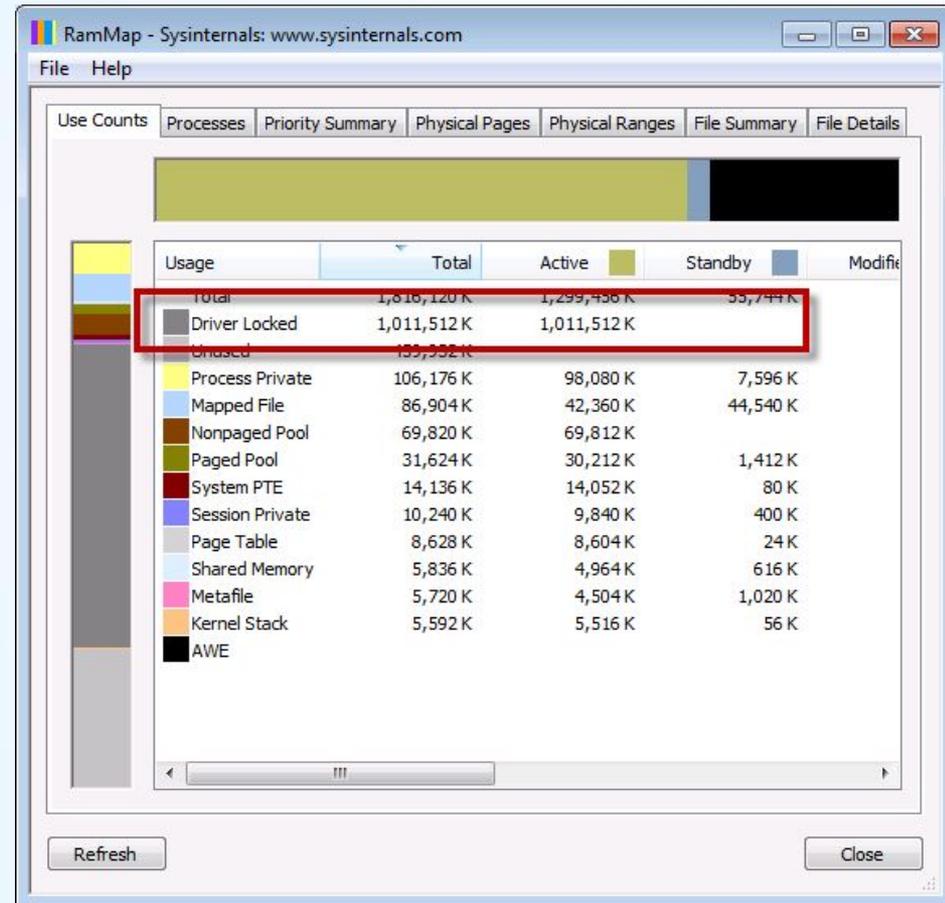
Добавление/удаление памяти

- Добавление памяти
 - Задействована гостевая ОС
 - Synthetic Memory Driver (VSP/VSC Pair)
 - Никакой эмуляции со стороны Hyper-V
 - Легкий и быстрый процесс
- Удаление памяти
 - Запрос на удаление памяти
 - Использование Ballooning
 - «Портит» показания task manager в гостевой ОС

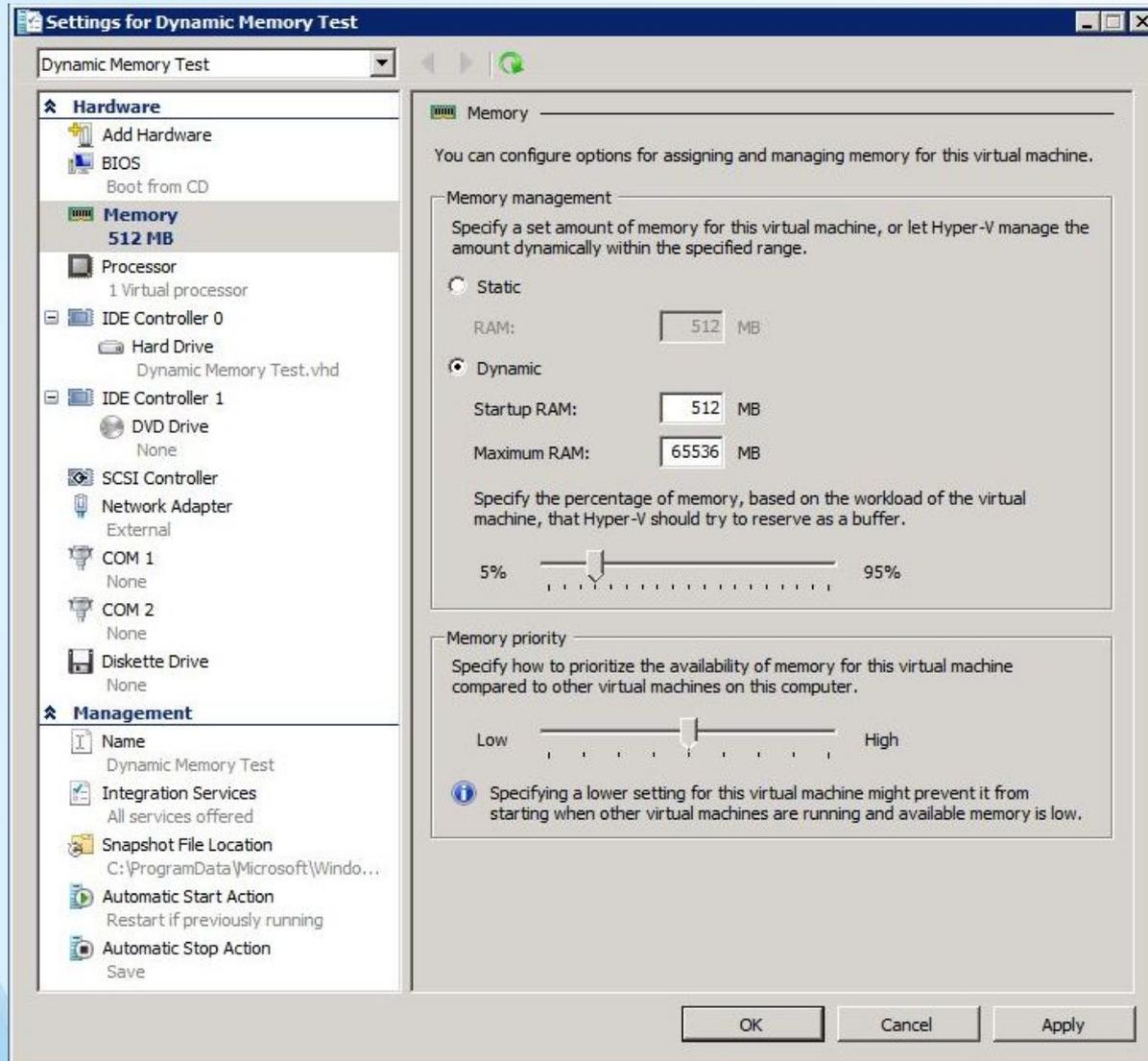


Уменьшение памяти

- Для уменьшения памяти используется balloon driver
 - «Отчуждению» подлежат только неиспользуемые области памяти
 - При «отчуждении» область памяти захватывается драйвером balloon и помечается как Driver Locked
 - При добавлении памяти драйвер может возвращать захваченные адресные пространства
- Таким образом, гостевая ОС «не замечает», что память «отнималась». Для гостевой ОС память как была, так и остается, лишь на время захваченная драйвером.



Dynamic Memory



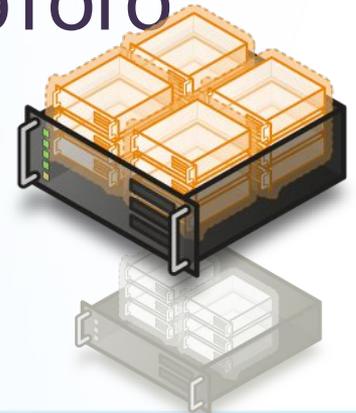
Требования

- Требования к хосту:
 - Windows Server 2008 R2 SP1
 - Microsoft Hyper-V Server 2008 R2 SP1
- Требования к гостевой ОС:
 - Windows Server 2003, 2008 & 2008 R2
 - Web, Standard, Enterprise и Datacenter Editions
 - 32-бита и 64-бита
 - Windows Vista and Windows 7
 - Enterprise и Ultimate Editions
 - 32-бита и 64-бита

Архитектура Dynamic Memory

Startup & Max

- Startup: достаточно памяти для старта VM
 - BIOS ничего не знает DM
 - Гостевая OS может ничего не знать о DM
 - Default: 512MB
- Max: не назначайте VM больше этого максимального значения
 - Default: 64GB



Pressure & Priority

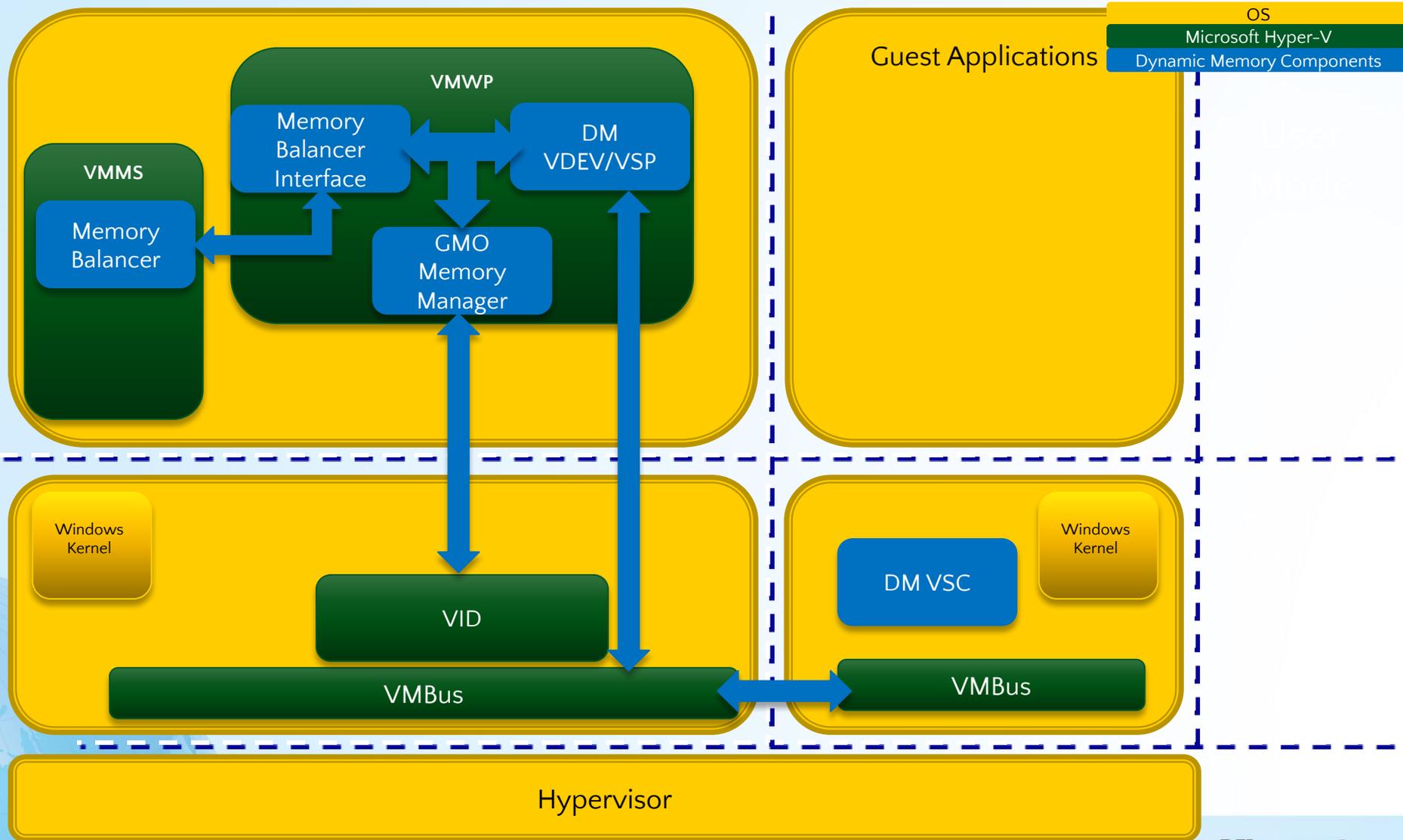
- Pressure – это концепт
 - Как много памяти у VM сейчас?
 - Сколько памяти VM требует?
 - Отношение и есть «pressure»
 - Работает с “committed memory”
- Priority: какая VM получает память первой
 - 1-10,000: default is 5,000
 - Более высокое значение более приоритетно

Memory Buffer

- Как много “свободной” памяти мы должны оставить для VM?
 - Позволяет реагировать на «пульсирующие» нагрузки в гостевой ОС
 - Может быть занята по кеш

“Я хотел бы сконфигурировать мои VM для того, чтобы у них было около -20% свободной памяти”

Архитектура Dynamic Memory



Dynamic Memory *demo*

Dynamic Memory APIs

- Доступен публично
 - [http://msdn.microsoft.com/en-us/library/cc136856\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/cc136856(VS.85).aspx)
 - Используются WMI-интерфейсы DMTF

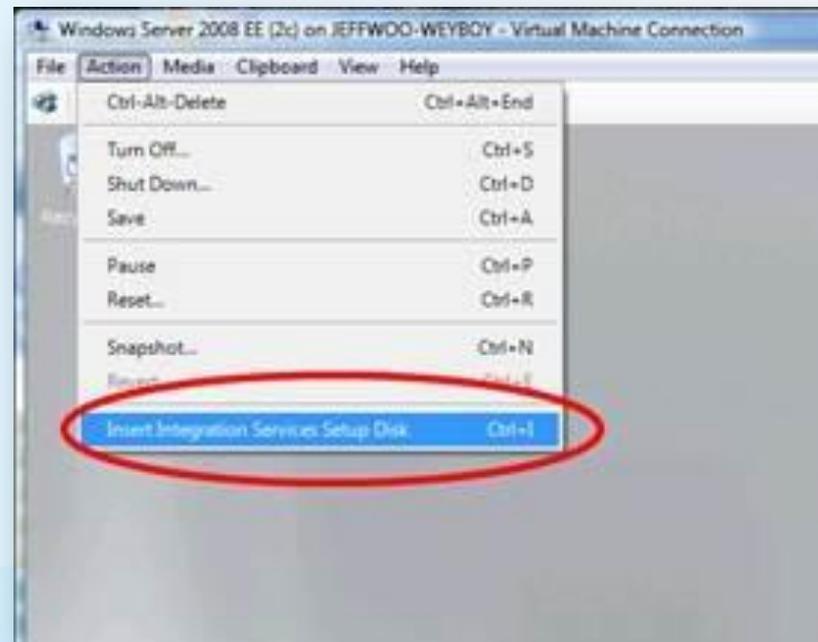
Hyper-V UI	Hyper-V WMI
Maximum Memory	Limit
Minimum Memory	Reservation
Startup Memory	VirtualQuantity
Priority	Weight
Dynamic Memory Buffer	TargetMemoryBuffer
Dynamic memory enabled	DynamicMemoryEnabled

Переход на R2 Service Pack 1

Что нужно сделать для успешного
использования Dynamic Memory?

3 шага...

- Убедиться, что VM готовы к обновлениям
- Обновить Хост
- Обновить Guest Integration Services



Заключение

- Memory Overcommit – маркетинговое пугало
- DM превращает память в такой же динамический ресурс, как и процессоры
- DM позволяет добиться большей консолидации без ущерба для производительности



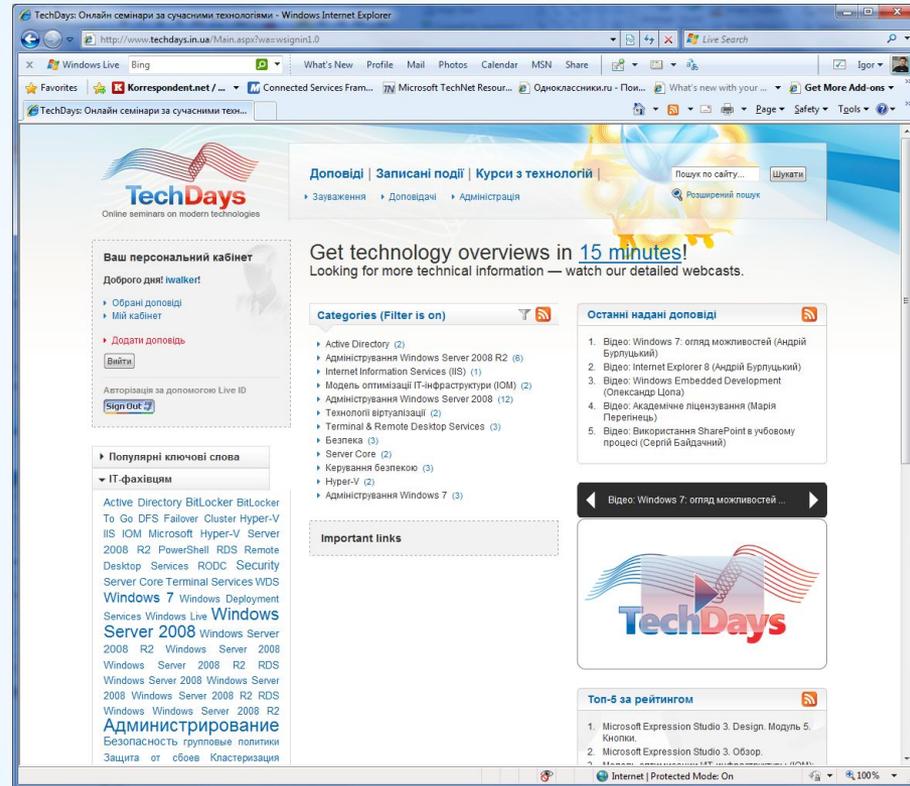
Microsoft® 
Kyiv Ukraine
2 - 3 December
SWIT 2010

Вопросы?

Microsoft®
Be what's next.™

www.techdays.in.ua

- <http://www.techdays.in.ua>
 - Новый информационный видеоресурс и портал
 - Содержит видеоматериалы с семинаров, курсов, вебкастов и т.п.
 - По различным инфраструктурным решениям
 - Используйте для обучения, справочной информации, знакомства с новыми технологиями
 - Пополняется ориентировочно несколько раз в неделю



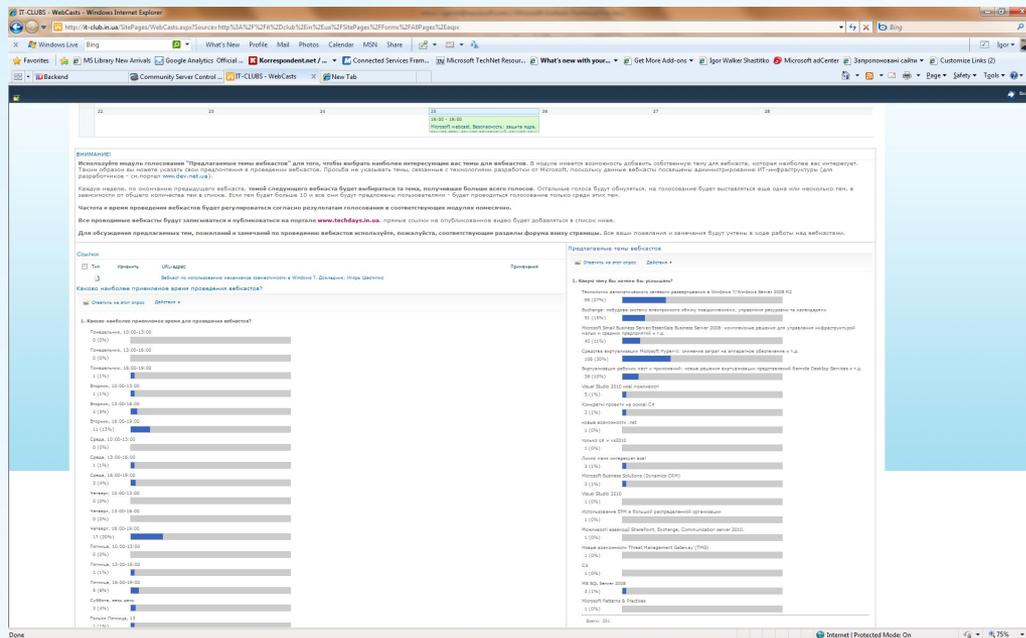
It-club.in.ua

Портал ИТ-клубов
Украины

www.it-club.in.ua

Создан и
поддерживается
энтузиастами
Киевского ИТ-клуба

В разделе «ВЕБКАСТЫ»
находится информация
о проведении вебкастов,
голосование,
обсуждение.



Основные цели портала

- Объединение ИТ-клубов Украины
- Информирование ИТ-специалистов о работе клубов и событиях
- Публикация информации об украинских блогерах
- Публикация статей и вебкастов
- Предоставление хостинга для сайтов ИТ-клубов и персональных технических блогов

Контакт: Сергей Бельский,
sergey.belskiy@hotmail.com



Microsoft[®]

© 2010 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Microsoft
Be what's next.™