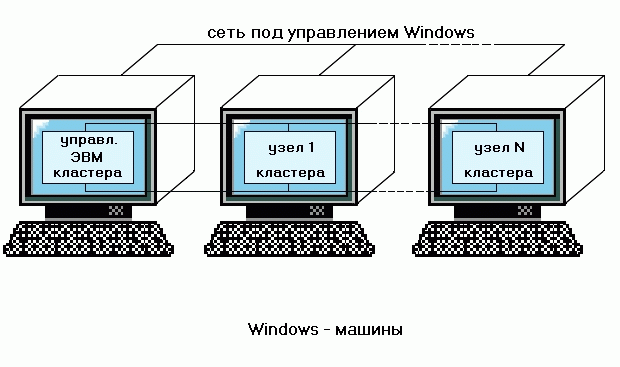
### Испытательная лаборатория проекта МВС.

### А. О. Лацис

# Руководство по установке МВС-900.

### Краткая справка об устройстве МВС-900

МВС-900 – это параллельная вычислительная система кластерной архитектуры, состоящая из одного хоста (управляющей машины) и нескольких (до 254) узлов. Хост и узлы – это виртуальные машины, каждая из которых выполняется на отдельной физической машине.



Общая структура виртуального кластера МВС-900.

Для реализации виртуальной машины в рамках физической используется монитор виртуальных машин VMware.

Конкретную конфигурацию виртуальной машины, которую требуется реализовать, VMware узнает из конфигурационного файла виртуальной машины. В этом файле указывается объем оперативной памяти, количество и объем виртуальных дисков, MAC- адреса виртуальных сетевых карт и прочие подобные сведения. Конфигурационный файл – текстовый, его можно редактировать обычным порядком. Конечно, программное обеспечение МВС-900 рассчитано на совершенно конкретную конфигурацию узлов и хоста.

Несколько слов о том, как именно VMware моделирует внешние устройства виртуальных машин на физическом оборудовании.

Проще всего обстоит дело с дисками. Каждая из виртуальных машин имеет хотя бы один виртуальный диск, которому в Windows физической машины соответствует один или несколько файлов, называемых файлами с образом диска. В состав дистрибутива входят файлы с образами дисков хоста и узла, на которых уже установлено все необходимое программное обеспечение кластера МВС-900.

Также каждая виртуальная машина оснащена несколькими виртуальными сетевыми картами. Отображение виртуальных сетевых карт на физическую сеть VMware выполняет следующим образом.

Физическая сетевая карта и все отображаемые на нее виртуальные сетевые карты логически просто объединяются виртуальным hub’ом, который и выходит на физический сетевой кабель.

Таким образом, на уровне MAC- адресов виртуальные сетевые карты оказываются в том же сегменте локальной сети, что и моделирующие их физические карты. Возможность сетевого взаимодействия виртуальных и физических сетевых карт, тем самым, не содержит никакой «виртуальной специфики», и зависит только от настройки ip- адресов. Все работает в точности так, как если бы упомянутый выше «виртуальный hub» существовал в действительности, а виртуальные сетевые карты были физическими и принадлежали отдельно стоящим физическим машинам. В частности, нет никакой разницы во взаимодействии виртуальной машины с той физической, на которой она установлена, и с любой другой физической или виртуальной машиной этого сегмента. Если ОС виртуальной машины настроит виртуальную сетевую карту на ip- адрес и маску из той же подсети, к которой относятся физические сетевые карты, все они смогут «разговаривать» друг с другом. Если же ОС нескольких виртуальных машин настроят свои сетевые карты на ip- адреса и маску другой подсети, эта подсеть будет существовать независимо от физической, отображаясь на то же оборудование. Виртуальные машины будут «разговаривать» друг с другом, не мешая физическим, и наоборот.

В МВС-900 используются оба варианта настройки ip- адресов. Так, хост и узлы связаны между собой двумя внутренними сетями, изолированными от внешнего мира, и, кроме того, хост имеет внешний сетевой интерфейс, позволяющий ему общаться с внешним миром, то есть с физическими машинами, а также с любым другим компьютером в Интернет, с которым связана локальная сеть физических машин.

На следующем рисунке изображена структура внутренних сетей МВС-900 (выход хоста во внешнюю сеть не показан):

**Сеть управления**

**Узел**

**N**

**Узел**

**1**

**Узел**

**2**

**Управля-**

**ющая**

##### ЭВМ

**Сеть коммуникаций**

Схема внутренних сетей МВС-900.

Конфигурирование МВС-900 при установке сводится, в основном, к конфигурированию сетевых интерфейсов узлов и хоста. Конфигурирование внешнего сетевого интерфейса хоста производится в последнюю очередь, с локальной консоли хоста, средствами Linux. Об этом см. ниже. Конфигурирование внутренних сетевых интерфейсов как хоста, так и узлов происходит по единой схеме, разработанной специально для МВС-900.

Схема состоит в следующем. Обычно результатом конфигурирования сети является запись на диск конфигурируемой машины заданной при конфигурировании информации (ip- адреса, сетевой маски, имени машины и т. п.). В МВС-900 так поступать было бы крайне неудобно – это означало бы, что при установке каждого узла с единого образа диска в этот образ диска приходилось бы вносить изменения, и потом сохранять не просто образ диска узла, а образ диска узла №1, №2 и так далее. Чтобы этого избежать, в МВС-900 ip- адреса во внутренних сетях и все сопутствующие настройки генерируются автоматически, по значениям MAC- адресов внутренних интерфейсов. MAC- адрес виртуальной сетевой карты задается в конфигурационном файле виртуальной машины. Как мы уже знаем, это совсем небольшой текстовый файл, поправить в котором «по месту» одну – две цифры много проще, чем вносить изменения в файл образа диска и затем хранить измененный образ.

Конкретная схема адресации во внутренних сетях устроена так.

VMware позволяет задавать в качестве MAC- адресов значения вида

00:50:56:00:X:Y

где X и Y выбираются произвольно. Именно эти поля в МВС-900 используются для задания полей ip- адресов.

В свою очередь, ip- адреса внутренних сетей МВС-900 имеют вид:

10.X.C.N для одной сети, и

10.X.C+32.N для второй, маски в обеих сетях – 255.255.224.0

Значения X и C выбираются системным администратором по его усмотрению, и должны быть одинаковыми для хоста и всех узлов. По умолчанию, X==100, C==5, и менять эти значения без особых на то оснований смысла нет.

Значение N – это номер узла. Для хоста N равно 1, для узлов лежит в диапазоне от 2 до 254 (не обязательно подряд).

В заданных в конфигурационном файле виртуальной машины MAC- адресах значения X, C и N кодируются следующим образом:

ethernet0.address = 00:50:56:00:X:N

ethernet1.address = 00:50:56:00:C:N

Следует учитывать, что MAC- адреса задаются в шестнадцатиричном виде.

Таким образом, в конфигурационном файле виртуальной машины хоста строки, задающие MAC- адреса, по умолчанию имеют вид:

ethernet0.address = 00:50:56:00:64:1

ethernet1.address = 00:50:56:00:05:1

и менять их нет оснований, если у Вас нет специальных претензий к выбранным значениям C и X.

В свою очередь, в конфигурационном файле виртуальной машины узла строки, задающие MAC- адреса, по умолчанию имеют вид:

ethernet0.address = 00:50:56:00:64:fe

ethernet1.address = 00:50:56:00:05:fe

то есть дистрибутивный узел имеет номер 254. Чтобы сделать узел, например, третьим, надо написать:

ethernet0.address = 00:50:56:00:64:3

ethernet1.address = 00:50:56:00:05:3

Если значения X и C (0x64 и 0x5) менялись для хоста, их надо поменять соответственно для всех узлов кластера.

Теперь мы знаем, что для установки комплекса МВС-900 требуется:

* установить на каждой из выбранных физических машин VMware,
* записать на каждую из машин файлы с образами дисков виртуальной машины, и соответствующие конфигурационные файлы,
* поправить в конфигурационных файлах несколько символов, чтобы машины получили правильные адреса во внутренних сетях,
* запустить каждую из виртуальных машин.

Пока мы, так или иначе, рассмотрели все пункты, кроме последнего. В самом деле, запускать хост и каждый из узлов вручную, физически обходя все физические машины и щелкая мышкой по соответствующим иконкам, причем всякий раз, когда требуется поработать на МВС-900, было бы довольно неудобно. Поэтому в МВС-900 используется, по мере возможности, автоматический запуск виртуальных машин в фоновом режиме.

Так, для узлов виртуальная машина запускается в фоновом режиме автоматически, при загрузке Windows, и работает, пока работает Windows. Никаких специальных действий по выключению узла МВС-900 не требуется – для выключения узла достаточно завершить работу Windows обычным порядком. При этом не гарантируется сохранность данных, записываемых на диск узла по мере его работы, но это и не требуется. МВС-900 устроена так, что все данные, которые должны сохраняться при выключении машин, хранятся на хосте. От узла лишь требуется сохранять на виртуальном диске данные в процессе его работы, от включения до выключения. При каждом включении узла его виртуальный диск имеет некоторый первоначальный вид – ровно такой, какой зафиксирован в дистрибутивном образе диска узла.

Чтобы обеспечить автоматический запуск узла при старте Windows, запуск виртуальной машины оформляется как сервис. Это делается однократно, при установке комплекса.

В случае хоста обеспечить полностью автоматические запуск и завершение работы виртуальной машины не представляется возможным, поскольку записываемые на виртуальный диск хоста во время работы данные должны сохраняться корректным образом в файле с образом диска при выключении хоста. Для этого требуется, как и в случае физической, «железной» машины, аккуратно завершить работу операционной системы, и, кроме того, аккуратно выключить виртуальное питание, а также аккуратно завершить работу самого приложения VMware. Словом, для завершения работы хоста требуются некоторые явно выполняемые вручную действия, поэтому хост все равно «нуждается в присмотре», и автоматический запуск его не слишком актуален. В общем случае, рекомендуется запускать его вручную, щелкнув дважды по иконке приложения VMware, выбрав соответствующую виртуальную машину и щелкнув “Power on”.

Общий обзор порядка установки и работы МВС-900 на этом завершен. Обзор этот касается лишь собственно установки комплекса и его общего взаимодействия с локальной сетью, на которую он ставится, но никак не работы на самой МВС-900. Краткое введение в порядок работы можно найти в документе howto.doc, а более подробное описание – в документации по МВС-1000, виртуальным вариантом которой и является МВС-900.

Ниже описывается порядок установки хоста и узла с дистрибутива.

### Порядок установки узла.

1. Однократно получить на физической машине администраторские права. Стать администратором. Здесь и далее имеется в виду локальный администратор, а не администратор домена.
2. Установить VMware. Дистрибутив – на этом диске. Место для установки лучше выбрать по умолчанию.
3. После установки VMware в панели управления, в разделе «Сеть и удаленный доступ», появятся два новых сетевых подключения, в названии которых присутствует слово “VMware”. Эти сетевые подключения следует выключить.
4. Создать директорию C:\mvs900. В действительности, здесь и далее имена создаваемых директорий и прочих системных объектов могут быть любыми, но в инструкции, для простоты изложения, используются конкретные значения.

Скопировать с дистрибутивного диска директорию tocopy в C:\mvs900. В командной строке Windows, сделать директорию C:\mvs900\tocopy текущей, и выполнить в ней команду распаковки образа виртуального диска:

gzip –d node.vmdk.gz

1. В winnt\system32 записать instsrv.exe и srvany.exe, если их там еще нет.
2. Выполнить в командной строке Windows:

instsrv startnode c:\winnt\system32\srvany.exe

1. Сделать текущей директорию дистрибутива, в которой находится файл parameters.reg, и выполнить в командной строке Windows:

start parameters.reg

При этом в реестре, в разделе

HKEY\_LOCAL\_MACHINES\SYSTEM\CurrentControlSet\Services\startnode

Образуется подраздел Parameters, в нем – следующие значения типа STRING:

Application= C:\ winnt\system32\cmd.exe

AppParameters= /d /c c:\mvs900\tocopy\startnode.bat

Если место для размещения дистрибутивной директории tocopy выбрано отличным от указанного выше, в п. 3, изменить значение второго параметра.

Замечание. В директории с:\mvs900\tocopy\node файла node.vmx нет. Это не ошибка – так и должно быть.

1. В панели управления найти меню управления сервисами, а в нем – только что созданный сервис startnode. Способ запуска сервису startnode указать «Автоматически», порядок запуска – от имени администратора. Временно установить режим разрешения вывода окон на Desktop.
2. Пустить вручную сервис startnode. Он должен запустить виртуальную машину. Поскольку для сервиса startnode вывод окон на Desktop пока разрешен, должно появиться окно VMware. При старте виртуальной машины появится одно или несколько небольших окошек с предложениями как-то ее оптимизировать, выполнить какие-то дополнительные действия и т. п. Во всех этих окошках имеется место для галочки «never show this hint again» (никогда больше не показывать эту подсказку). Эту галочку везде надо проставить. Когда виртуальная машина стартует, поднять ей до максимума Local priority и Global priority. Затем выключить виртуальное питание, завершить VMware, остановить сервис startnode и снова пустить. Повторять эту процедуру (завершение и повторный запуск), пока старт виртуальной машины не станет «чистым», безо всяких вопросов в виде дополнительных окошек. Еще раз все завершить, не пуская повторно.
3. Стать администратором. Запретить сервису startnode вывод окон на Desktop, способ запуска сервиса сделать автоматическим, перевызвать Windows и, не выполняя logon, убедиться (по звуковому сигналу), что виртуальная машина стартовала.
4. Зайти как администратор в последний раз, убрать с Desktop иконку VMware, чтобы какой-нибудь пользователь из любопытства не запустил ее и не поменял жизненно важных настроек.

Замечание. В этом состоянии система пригодна для копирования в виде образа системного диска Windows. Дальнейшая настройка касается лишь обеспечения индивидуальности узлов (присвоения им уникальных номеров).

11.Присвоить узлу уникальный номер. Номер узла должен лежать в диапазоне от 2 до 254, включительно, и быть уникальным в кластере. MAC-адреса интерфейсов узла задаются в файле C:\mvs900\tocopy\node.saved\node.vmx, в строках ethernet0.address и ethernet1.address, соответственно.

Упомянутые строки должны иметь вид:

ethernet0.address = 00:50:56:00:X:N

ethernet1.address = 00:50:56:00:C:N

например

ethernet0.address = 00:50:56:00:64:N

ethernet1.address = 00:50:56:00:05:N

где N – номер узла в кластере, X (0x64==100) – номер куста кластеров, он же второй байт ip-адреса всех машин в кусте, C (0x5==5) – номер кластера, он же младшие 5 бит третьего байта (см. выше). Изменения вступают в силу при перезапуске виртуальной машины (придется опять перегрузить Windows).

### Порядок установки хоста.

1. Однократно получить на физической машине администраторские права. Стать администратором. Здесь и далее имеется в виду локальный администратор, а не администратор домена.
2. Установить VMware. Дистрибутив – на этом диске. Место для установки лучше выбрать по умолчанию.
3. Создать директорию C:\mvs900. В действительности, здесь и далее имена создаваемых директорий и прочих системных объектов могут быть любыми, но в инструкции, для простоты изложения, используются конкретные значения.

Скопировать с дистрибутивного диска директорию tocopy в C:\mvs900.

Если дистрибутивных дисков для хоста несколько, скопировать со всех. Распаковать все файлы с именами вида host\*.vmdk.gz при помощи gzip: gzip –d <имя файла>. Выполнить сетевое конфигурирование хоста во внутренних сетях, отредактировав файл C:\mvs900\tocopy\host.vmx так, как это объяснено выше.

1. Запустить VMware. Выбрать виртуальную машину C:\mvs900\tocopy\host, и пустить ее. Переключение клавиатуры между виртуальной и физической машинами: для перехода на виртуальную, щелкнуть мышью по внутренней области ее окна. Для возврата в физическую – Ctrl+Alt.
2. Войти в Linux как root. Пароль – virtualcluster. Настроить сетевой адрес во внешней сети командой netconfig –d eth2. Сменить пароль root’а. Замечание: команда netconfig в указанном выше виде характерна для RedHat, в то время как Вы уже заметили, что на управляющей виртуальной машине установлен Slackware Linux. Это не ошибка – команда netconfig из Slackware доработана так, чтобы она воспринимала указанные выше аргументы.
3. Правильная команда регистрации пользователя на кластере лежит в /common/howtoaddauser. Регистрировать надо только на хосте. Документация – в /common/runmvs/doc. Файлы переписывать при помощи scp, например, при помощи pscp.exe, или Winscp.
4. Выключение хоста: войти как root, выполнить команду:

shutdown –h now

После завершения Linux виртуальное питание выключится автоматически. Завершить приложение VMware.

### Замена образа диска или файла конфигурации виртуальной машины узла.

#### При необходимости обновить версию узла (например, заменить файл с образом диска) следует учитывать, что остановка сервиса “startnode” не останавливает виртуальную машину. «Выдергивание» открытого файла из-под работающей виртуальной машины не желательно, точнее, просто невозможно. Поступать следует так.

* в панели управления отключить сервис “startnode”,
* перезагрузить Windows,
* заменить файлы, относящиеся к виртуальной машине,
* включить сервис “startnode” в режиме автоматического запуска,
* перезагрузить Windows.