WHITE PAPER

ПОДКЛЮЧЕНИЕ РЕСУРСОВ К СЕРВЕРНОЙ ОС: ЗОНИРОВАНИЕ FC SAN

YADRO TATLIN

О WHITE PAPERS

Данный документ описывает архитектурные принципы и подходы, а также рекомендации по их применению при построении корпоративных решений на основе продуктов производства компании YADRO.

Документ предназначен для широкой аудитории, но также может включать материалы, требующие специальной профессиональной подготовки, ориентированные на системных архитекторов, администраторов, сервисных инженеров и других IT-специалистов, участвующих в планировании, проектировании и внедрении комплексных решений.

! Представленная информация основана на результатах исследований, а также экспертизе специалистов YADRO. Документ носит не предписывающий, а информационный характер, целью которого является информирование получателей документа о предмете произведенного исследования и о его результатах. При этом конкретные результаты и эффективность конечных решений зависят от множества факторов, которые могут различаться в зависимости от особенностей инфраструктуры и условий эксплуатации. Ни при каких обстоятельствах YADRO не несет ответственность за последствия применения получателем документа рекомендаций, инструкций, подходов и решений, указанных в документе.

Запрещается копировать, воспроизводить, изменять, публиковать, передавать, распространять, продавать информацию (часть информации), содержащуюся в документе, или создавать производные работы на основе содержания документа без предварительного письменного разрешения YADRO.

Любая информация или утверждения относительно возможностей, емкости, производительности или других характеристик продукции, указанной в документе, предоставляются исключительно на условиях «как есть» и приведена исключительно в целях описания произведенного исследования. YADRO оставляет за собой право вносить изменения/обновлять информация в документе без уведомления.

«YADRO» является зарегистрированным товарным знаком ООО «КНС ГРУПП». Другие названия продуктов и компаний, упомянутые здесь, могут быть товарными знаками или торговыми наименованиями их соответствующих владельцев.

Актуальные версии документации публикуются на информационном ресурсе YADRO по адресу service.yadro.com.

ОГЛАВЛЕНИЕ

1. ОБ ЭКСПЕРТАХ 4
2. ВВЕДЕНИЕ 5
3. ТИПЫ ЗОНИРОВАНИЯ 6
4. РАЗРАБОТКА ПЛАНА ЗОНИРОВАНИЯ 6
5. ЗОНИРОВАНИЕ С ОДНИМ ИНИЦИАТОРОМ (SINGLE INITIATOR ZONING) 7
6. PEER ZONING – НОВЫЙ ПОДХОД К ЗОНИРОВАНИЮ 8
7. ПРИМЕРЫ НАСТРОЙКИ ЗОНИРОВАНИЯ 10
8. ПРИМЕР НАСТРОЕК ЗОНИРОВАНИЯ 11
9. ПРАКТИЧЕСКИЕ АСПЕКТЫ ИСПОЛЬЗОВАНИЯ МЕХАНИЗМОВ FC-ЗОНИРОВАНИЯ 12

ОБ ЭКСПЕРТАХ

Дмитрий Боровской — специалист с более чем 30-летним опытом работы в сфере ИТ, включая работу с системами хранения данных различных вендоров (EMC, Dell, Hitachi) и отечественных производителей (YADRO). В YADRO занимает должность ведущего специалиста отдела Мультивендорной сервисной экспертизы. Специализируется на администрировании Linux, конфигурировании и поддержке систем хранения, а также интеграции продуктов TATLIN в инфраструктуру заказчиков. Эксперт в области сетей хранения и передачи данных (SAN, IPSAN).

ВВЕДЕНИЕ

С помощью зонирования администратор SAN-сети может разбить фабрику на зоны, объединяющие взаимодействующие между собой устройства, и применить лучшие в своем классе механизмы блокировки связи между устройствами из разных зон. Зоны могут перекрываться и поддерживать как очень простые, так и сложные схемы передачи трафика. Зонирование можно рассматривать как комбинирование лучших аспектов технологии VLAN с мощными функциями безопасности, обычно используемыми только в межсетевых экранах корпоративного класса.

Зоны удобны для создания барьеров между разными операционными средами, создания внутри фабрики нескольких функциональных групп или тестовой среды и/или областей, изолированных от остальной фабрики. Если в фабрике используются зоны, то никакой сервер или устройство хранения не смогут использовать фабрику до тех пор, пока они явно не включены хотя бы в одну зону авторизованным администратором фабрики.

Зонирование — это инструмент уровня фабрики, и конфигурации зон автоматически передаются на все ее коммутаторы. Фабрика обеспечивает максимум резервирования и надежности, поскольку каждый коммутатор хранит информацию о зонах и может ее передать новому коммутатору, добавленному в фабрику.

ВЕРСИИ ИСПОЛЬЗУЕМОГО ПО	
TATLIN.UNIFIED Gen1/Gen2	2.7–3.1.x
FabricOS	8.1.x–9.2.x

ТИПЫ ЗОНИРОВАНИЯ

Самое простое зонирование — на базе портов (hard zoning), то есть физические порты коммутатора распределяются между разными зонами. Например, задается правило: только устройства, подключенные к порту 1 коммутатора 1, могут «разговаривать» с устройствами на порте 2 коммутатора 3.

В настоящее время предпочтительным методом стало зонирование на базе WWN, которое ограничивает доступ по WWN устройств вместо идентификатора порта (soft zoning). Это более гибкий метод, поскольку любые узлы сети всегда остаются в назначенной им зоне. Однако у него есть один недостаток — если заменить устройство или FC HBA, то у него изменится WWN, поэтому потребуется вносить обновление в конфигурацию зон. Разумеется, переключение устройства от одного порта к другому при использовании зонирования по WWN выполняется проще, поэтому этот недостаток не считается принципиальным.

РАЗРАБОТКА ПЛАНА ЗОНИРОВАНИЯ

Хотя текущим управлением зонированием обычно занимается администратор SAN, архитектор SAN должен продумать, как построить зоны, поскольку от этого зависит, какие потоки данных между устройствами будут разрешены. Также архитектор еще на стадии проектирования должен проанализировать характер потоков данных с точки зрения производительности, что упростит разработку плана зон.

Здесь важен творческий подход, поскольку не может быть только одной «правильной» конфигурации зон для данной сети. Далее описывается один из распространенных подходов.

Сначала необходимо составить перечень всех серверов и систем хранения данных (СХД), которые нужно объединить в зоны (обычно к этому моменту уже имеется перечень подключенных к SAN устройств). В зависимости от привычных для архитектора SAN утилит данный перечень может быть представлен в виде электронной таблицы, текстового файла или другом формате и должен включать имя, физическое расположение, WWN, идентификатор порта и функцию каждого устройства.

Затем необходимо оценить потребности каждого сервера в хранении, и если ему выделено отдельное устройство, указать это в списке, а если нет, определить, сколько устройств хранения и какого типа потребуется серверу. Далее необходимо найти ближайшую СХД, которая соответствует требованиям сервера и выделить серверу емкость из этого дискового массива. В списке необходимо указать, что серверу требуется доступ к этому массиву. Необходимо предусмотреть сценарии переключения при сбоях, несколько портов под управление программного обеспечения multipathing и доступ к серверам резервного копирования и лентам.

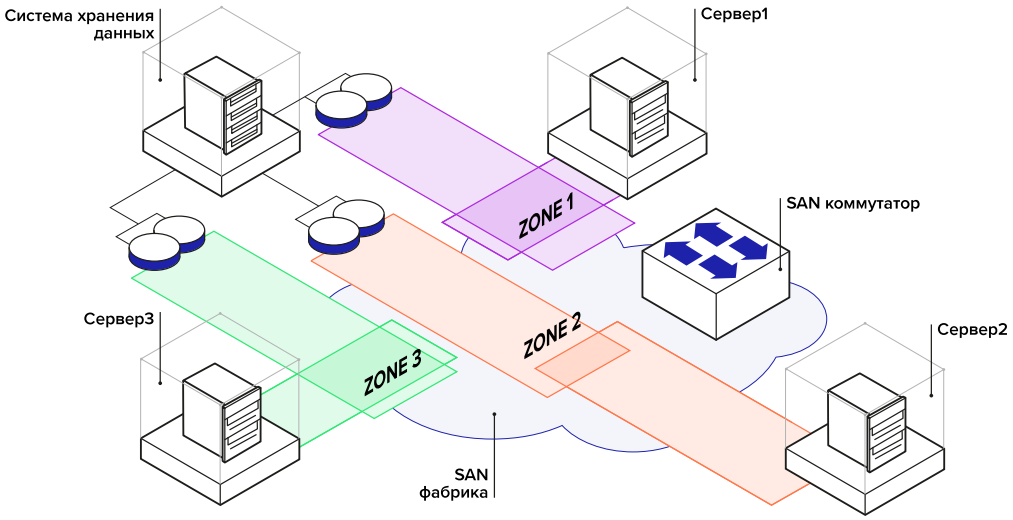
После того как у архитектора появится понимание, какому порту сервера нужен доступ к какому порту устройства хранения, необходимо записать эти сведения.

Например, архитектор может сконфигурировать одну зону из одного порта СХД и всех серверов, которые обращаются к нему. При этом один сервер может обращаться к другому серверу из этой же зоны, что в SAN обычно не требуется и даже нежелательно. Поэтому лучше иметь по одной зоне на пару устройств и использовать перекрывание зон. В перекрывающихся зонах адаптеры серверов совместно используют один или несколько портов СХД, но сами адаптеры изолированы друг от друга. В идеале фабрика состоит из большого числа зон с парами портов или WWN и определяет связь между парой «инициатор – таргет» (Initiator – Target). Такие зоны иногда называют point-to-point. Если инициатор сконфигурирован на доступ к нескольким портам устройств хранения данных, следует включить его в несколько зон вместо того, чтобы объединять несколько портов в одну зону устройств хранения, тогда будет возможен только тот доступ, который необходим.

ЗОНИРОВАНИЕ С ОДНИМ ИНИЦИАТОРОМ (SINGLE INITIATOR ZONING)

В очень больших сетях применение зон point-to-point не всегда удобно, поскольку оно означает создание большого числа зон и необходимость многочисленных изменений в конфигурации фабрики при добавлении и перемещении устройств. Более популярный метод — зоны с одним инициатором, при котором достигается баланс между необходимостью ограничить число зон и упрощением управления. Для каждого инициатора в фабрике создается отдельная зона, которая также содержит все порты систем хранения для этого инициатора.

Кроме обеспечения повышенной безопасности, данный подход помогает оптимизировать сервисы фабрики. Зоны оптимизируют сервисы фабрики, такие как распространение уведомлений (RSCN) и ответ сервера имен и ограничивают обнаружение «лишних» устройств. Корректная конфигурация зон необходима для больших сетей (даже если для них безопасность некритична), и методы point-to-point и «одного инициатора» являются оптимальными.



Зонирование с одним инициатором

Архитектору также необходимо разработать удобную систему именования зон. Для улучшения масштабируемости и управляемости лучше всего использовать короткие и понятные имена. Трудно придумать короткое, но в то же время понятное имя, поэтому рекомендуется сокращать имена так, чтобы они не потеряли своего значения.

Зоны должны содержать только алиасы (alias) устройств, но не номера портов или WWN, для того, чтобы можно было определить устройство только один раз и использовать его в нескольких зонах без переопределений. Если устройство вышло из строя, было перемещено или модернизировано, потребуется изменить только WWN, привязанный к его алиасу, после чего это изменение отразится на всех зонах, в которых он был прописан.

Например, читаемое и понятно название алиаса содержит имя сервера, номер контроллера и порта. Или же имя СХД, номер контроллера и номер порта контроллера: `Server1-hba1-port1`. Плохо читаемый пример: `CK000236587_Left` — использован серийный номер устройства, который труднее запомнить, контроллер указан как left, что затрудняет определение из-за возможной зеркальности, отсутствует указание порта.

Зонирование стоит использовать в комбинации с другими методами, такими как LUN mapping, LUN-masking.

PEER ZONING — НОВЫЙ ПОДХОД К ЗОНИРОВАНИЮ

В операционной системе Fabric OS, начиная с версии 7.4, была разработана достаточно интересная функциональность.

До версии Fabric OS 7.3 включительно рекомендуемым и наиболее распространенным методом зонирования оставалось зонирование с одним инициатором (single initiator zoning).

При таком методе зонирования для каждого инициатора в фабрике создается отдельная зона, включающая все порты систем хранения, с которыми взаимодействует данный инициатор. Этот метод считался наиболее популярным в течение долгого времени, так как позволял достичь баланса между числом сконфигурированных в фабрике зон и простотой управления.

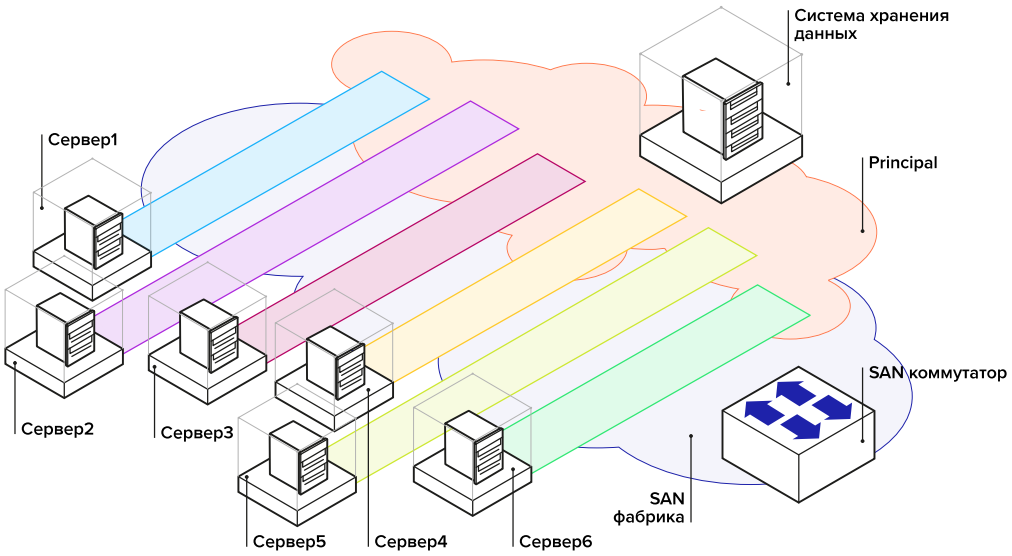
Однако с ростом размеров фабрик и числа подключенных к фабрикам устройств данный метод перестал соответствовать требованиям масштабируемости и управляемости. Количество сконфигурированных зон стремительно увеличивается, что усложняет их администрирование. В каждом коммутаторе сети SAN имеется специальная база данных (zone database), в которой хранится информация о зонировании фабрики. Поскольку объем памяти, выделенный для этой базы данных, ограничен, при увеличении числа зон может возникнуть ситуация, когда база данных достигнет максимального размера, что сделает дальнейшее управление зонами невозможным.

Для упрощения конфигурирования зон и сокращения записей в базе данных стандартами протоколов Fibre Channel следующего поколения (FC-SW-6 и FC- GS-7) предусмотрен новый тип зонирования — Peer Zoning. Хотя этот стандарт является обязательным для оборудования Gen 6, начиная с версии FOS 7.4 он стал доступен и для существующих устройств Gen 5.

При конфигурации зонирования методами Peer Zoning в базе данных создаются два типа участников — principal member и peer member.

Данная конфигурация позволяет взаимодействовать между собой любому principal member и любому peer member, входящим в одну peer zone. Но любые два peer member не могут взаимодействовать между собой. Также не могут взаимодействовать между собой любые два principal member.

Для типичной конфигурации, количество инициаторов (серверов) значительно превышает количество target (СХД). Поэтому инициаторы (серверы) конфигурируются как peer members, а порты СХД — как principal members. В одну зону можно включить необходимое количество как principal members, так и peer members.



Зонирование Peer Zoning: 8 Peer members + 1 Principal member

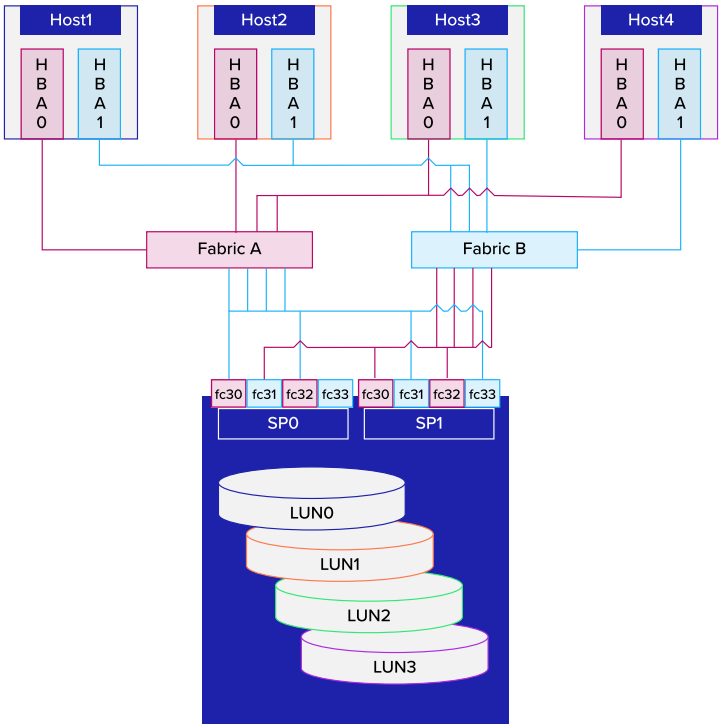
Преимущества конфигурации:

- Значительно упрощается конфигурация зонирования за счет сокращения количества зон по сравнению с традиционными методами зонирования;
- сокращается количество широковещательных сообщений в фабрике (RSCN distribution);
- эффективно используется база данных зонирования, сокращается нагрузка на внутренние процессоры коммутатора.

В примере выше, в случае использования Peer Zoning, конфигурация сокращается до одной зоны, в которую входят 8 peer members и один principal member.

Применяемый в Fabric OS 7.4 метод конфигурирования и использования Peer Zoning полностью соответствует новым стандартам FC-SW-6 и FC-GS-7. При этом зонная конфигурация фабрики может содержать как традиционные зоны, так и новые зоны с использованием Peer Zoning одновременно. Peer Zoning также может использоваться при маршрутизации между фабриками (Fibre Channel Routing).

ПРИМЕРЫ НАСТРОЙКИ ЗОНИРОВАНИЯ



- На рисунке выше сеть SAN состоит из двух фабрик: Fabric A и Fabric B.
- К Fabric A подключены FC порты HBA0 серверов host1... host4 и FC порты fc30, fc32 контроллера SP0 СХД и порты fc31, fc33 контроллера SP1 СХД.
- К Fabric B подключены FC порты HBA1 серверов host1... host4 и FC порты fc30, fc32 контроллера SP1 СХД и порты fc31, fc33 контроллера SP0 СХД.

!

Данная схема подключения является требованием к FC коммутации СХД Tatlin.Unified в силу его архитектурных особенностей. Ввиду того, что Tatlin.Unified использует Active-Active архитектуру доступа, FC порты на каждом контроллере объединены в виртуальную пару. При маппировании ресурса допускается выбрать только один порт, например fc30. При этом ресурс будет выдан для серверного доступа по двум путям через оба контроллера. Из этой особенности архитектуры вытекает требование к отказоустойчивости подключения — подключать одноименные порты контроллеров в разные фабрики.

На приведенной схеме порты fc30 и fc32 контроллера sp0 и fc31, fc33 контроллера sp1 подключены в Fabric A, а порты fc31,fc33 контроллера sp0 и порты fc30,fc32 контроллера sp1 подключены в Fabric B.

ПРИМЕР НАСТРОЕК ЗОНИРОВАНИЯ

Создаем алиасы

FABRIC A

Серверы:

```
alicreate "host1_p0", "21:00:f4:c7:aa:00:33:3a"
alicreate "host2_p0", "21:00:f4:c7:aa:00:2a:4c"
alicreate "host3_p0", "21:00:f4:c7:aa:00:4c:24"
alicreate "host4_p0", "21:00:f4:c7:aa:00:6f:78"
```

Порты массива:

```
alicreate "TU_sp0_fc30", "10:00:14:52:90:00:fb:18"
alicreate "TU_sp0_fc32", "10:00:14:52:90:00:fb:20"
alicreate "TU_sp1_fc31", "10:00:14:52:90:00:fb:99"
alicreate "TU_sp1_fc33", "10:00:14:52:90:00:fb:a1"
```

FABRIC B

Серверы:

```
alicreate "host1_p1", "21:00:f4:c7:aa:00:33:3b"
alicreate "host2_p1", "21:00:f4:c7:aa:00:2a:4d"
alicreate "host3_p1", "21:00:f4:c7:aa:00:4c:25"
alicreate "host4_p1", "21:00:f4:c7:aa:00:6f:79"
```

Порты массива:

```
alicreate "TU_sp0_fc31", "10:00:14:52:90:00:fb:19"
alicreate "TU_sp0_fc33", "10:00:14:52:90:00:fb:21
alicreate "TU_sp1_fc30", "10:00:14:52:90:00:fb:98"
alicreate "TU_sp1_fc32", "10:00:14:52:90:00:fb:a0"
```

One-to-one zoning (один инициатор, один таргет)

FABRIC A

```
zonecreate "h1p0_TU_sp0_fc30_z", "host1_p0; TU_sp0_fc30"
zonecreate "h1p0_TU_sp0_fc32_z", "host1_p0; TU_sp0_fc32"
zonecreate "h1p0_TU_sp1_fc31_z", "host1_p0; TU_sp1_fc31"
zonecreate "h1p0_TU_sp1_fc33_z", "host1_p0; TU_sp1_fc33"

zonecreate "h2p0_TU_sp0_fc30_z", "host2_p0; TU_sp0_fc30"
zonecreate "h2p0_TU_sp0_fc32_z", "host2_p0; TU_sp0_fc32"
zonecreate "h2p0_TU_sp1_fc31_z", "host2_p0; TU_sp1_fc31"
zonecreate "h2p0_TU_sp1_fc33_z", "host2_p0; TU_sp1_fc33"

zonecreate "h3p0_TU_sp0_fc30_z", "host3_p0; TU_sp0_fc30"
zonecreate "h3p0_TU_sp0_fc32_z", "host3_p0; TU_sp0_fc32"
zonecreate "h3p0_TU_sp1_fc31_z", "host3_p0; TU_sp1_fc31"
zonecreate "h3p0_TU_sp1_fc33_z", "host3_p0; TU_sp1_fc33"

zonecreate "h4p0_TU_sp0_fc30_z", "host4_p0; TU_sp0_fc30"
zonecreate "h4p0_TU_sp0_fc32_z", "host4_p0; TU_sp0_fc32"
zonecreate "h4p0_TU_sp1_fc31_z", "host4_p0; TU_sp1_fc31"
zonecreate "h4p0_TU_sp1_fc33_z", "host4_p0; TU_sp1_fc33"
```

FABRIC B

```
zonecreate "h1p1_TU_sp0_fc31_z", "host1_p1; TU_sp0_fc31"
zonecreate "h1p1_TU_sp0_fc33_z", "host1_p1; TU_sp0_fc33"
zonecreate "h1p1_TU_sp1_fc30_z", "host1_p1; TU_sp1_fc30"
zonecreate "h1p1_TU_sp1_fc32_z", "host1_p1; TU_sp1_fc32"

zonecreate "h2p1_TU_sp0_fc31_z", "host2_p1; TU_sp0_fc31"
zonecreate "h2p1_TU_sp0_fc33_z", "host2_p1; TU_sp0_fc33"
zonecreate "h2p1_TU_sp1_fc30_z", "host2_p1; TU_sp1_fc30"
zonecreate "h2p1_TU_sp1_fc32_z", "host2_p1; TU_sp1_fc32"

zonecreate "h3p1_TU_sp0_fc31_z", "host3_p1; TU_sp0_fc31"
zonecreate "h3p1_TU_sp0_fc33_z", "host3_p1; TU_sp0_fc33"
zonecreate "h3p1_TU_sp1_fc30_z", "host3_p1; TU_sp1_fc30"
zonecreate "h3p1_TU_sp1_fc32_z", "host3_p1; TU_sp1_fc32"

zonecreate "h4p1_TU_sp0_fc31_z", "host4_p1; TU_sp0_fc31"
zonecreate "h4p1_TU_sp0_fc33_z", "host4_p1; TU_sp0_fc33"
zonecreate "h4p1_TU_sp1_fc30_z", "host4_p1; TU_sp1_fc30"
zonecreate "h4p1_TU_sp1_fc32_z", "host4_p1; TU_sp1_fc32"
```

Single initiator zoning (один инициатор, несколько таргетов)

FABRIC A

```
zonecreate "h1_TU_z",
  "host1_p0; TU_sp0_fc30; TU_sp0_fc32; TU_sp1_fc31; TU_sp1_fc33"

zonecreate "h2_TU_z",
  "host2_p0; TU_sp0_fc30; TU_sp0_fc32; TU_sp1_fc31; TU_sp1_fc33"

zonecreate "h3_TU_z",
  "host3_p0; TU_sp0_fc30; TU_sp0_fc32; TU_sp1_fc31; TU_sp1_fc33"

zonecreate "h4_TU_z",
  "host4_p0; TU_sp0_fc30; TU_sp0_fc32; TU_sp1_fc31; TU_sp1_fc33"
```

FABRIC B

```
zonecreate "h1_TU_z",
  "host1_p1; TU_sp0_fc31; TU_sp0_fc33; TU_sp1_fc30; TU_sp1_fc32"

zonecreate "h2_TU_z",
  "host2_p1; TU_sp0_fc31; TU_sp0_fc33; TU_sp1_fc30; TU_sp1_fc32"

zonecreate "h3_TU_z",
  "host3_p1; TU_sp0_fc31; TU_sp0_fc33; TU_sp1_fc30; TU_sp1_fc32"

zonecreate "h4_TU_z",
  "host4_p1; TU_sp0_fc31; TU_sp0_fc33; TU_sp1_fc30; TU_sp1_fc32"
```

Peer Zoning

FABRIC A	FABRIC B
`zonecreate --peerzone "TU_z" -principal "TU_sp0_fc30; TU_sp0_fc32; TU_sp1_fc31; TU_sp1_fc33" -members "host1_p0; host2_p0; host3_p0; host4_p0"`	`zonecreate --peerzone "TU_z" -principal "TU_sp0_fc31; TU_sp0_fc33; TU_sp1_fc30; TU_sp1_fc32" -members "host1_p1; host2_p1; host3_p1; host4_p1"`

ПРАКТИЧЕСКИЕ АСПЕКТЫ ИСПОЛЬЗОВАНИЯ МЕХАНИЗМОВ FC-ЗОНИРОВАНИЯ

Использование зонирования, как логического разграничения коммутатора на изолированные и защищенные сегменты передачи данных по принципу привязки к СХД, серверу или приложениям, обеспечивает устойчивую работу всей FC-инфраструктуры, не позволяя, например, шторму событий-запросов от одного инициатора замедлить работу всей FC фабрики. При использовании зон все последствия такого шторма останутся в пределах той зоны, в которую добавлен неисправный инициатор. Использование зонирования в фабрике существенно облегчает управление и мониторинг фабрики, позволяя проводить операции в конкретной зоне без какого-либо влияния на остальное окружение.

Зоны настоятельно рекомендуется создавать на основе WWPN подключенных устройств, избегая использования для зонирования номеров физических портов. Для WWPN устройств необходимо использовать «говорящие» имена — алиасы (aliases). Это значительно упрощает администрирование и идентификацию устройств при добавлении их в несколько зон. Выбор методов зонирования зависит от размеров фабрики, количества устройств и предпочтений администратора. Как следует из вышеизложенного, при использовании One-to-One Zoning в каждой фабрике необходимо создать по 16 зон, при использовании Single Initiator Zoning — по 4 зоны. Использование Peer Zoning предполагает меньшие административные затраты, упрощает управление фабрикой, сохраняя все преимущества soft zoning. Peer Zoning является рекомендуемой конфигурацией для больших фабрик с множеством портов.

При применении изменений в конфигурации необходимо выполнять и активацию, и сохранение обновленной конфигурации, чтобы избежать возможных потерь изменений при перезагрузке коммутатора.