

Hitachi

TagmaStore USP:

Взгляд изнутри

7 сентября 2004 г. компания Hitachi Data Systems представила свою новую систему high-end класса – Hitachi TagmaStore™ Universal Storage Platform (USP). Это, безусловно, самое крупное объявление HDS за последние несколько лет, как по параметрам системы, так и по значению для рынка систем хранения данных.

Введение

Новая система обладает такими показателями: производительность около 2 000 000 IOPS¹, пропускная способность – 81 Гбайт/с и масштабируемость – свыше 300 Тбайт, для которых понятия “грандиозный” и “фантастический” не являются преувеличением.

Однако основная идея TagmaStore – предоставить потребителю универсальную платформу для консолидации всех ресурсов хранения – тех, что уже есть, и тех, которые

только будут. Отличие же данной реализации этой идеи заключается в том, что, подключая к, по существу, интеллектуальному дисковому массиву любую “голую железку”, вы бесплатно распространяете на нее возможности функционала USP, при этом обеспечивая сквозное управление всеми ресурсами. Таким образом вы сможете “упорядочить” (“tagга” в переводе с греческого – “порядок”) более 32 Пбайт(!) внешней дисковой памяти.

В функциональности TagmaStore нашли отражение основные тенденции, активно развиваемые в отрасли последние 3 года. “Тремя китами” USP являются: виртуализация хранения на основе консолидации ресурсов в общий пул, разделение общего пула на независимые по применению/использованию домены/разделы и расширенная по возможностям репликация данных.

Что нового в TagmaStore?

TagmaStore является дальнейшим эволюционным развитием серии Lightning 9900V (модели 9970V и 9980V), объявленной еще в июне 2000 г. и успешно продаваемой сейчас. Рассмотрим более подробно концептуальные нововведения, отличающие TagmaStore USP от предшественников.

Первое и наиболее крупное из новшеств – то, что делает TagmaStore™ универсальной платформой хранения данных (Universal Storage Platform – USP) – возможность объединения средствами TagmaStore™ внутренних и внешних по отношению к системе гетерогенных ресурсов хранения для совместного использования множеством серверов и мэйнфреймов (рис. 1). Для подключаемых внешних ресурсов хранения, которые могут иметь различную скорость доступа к данным (диски – FC/SAS/SATA, ленты), разный уровень надежности и доступности, TagmaStore USP обеспечивает единообразное представление, единую функциональность, необходимый уровень доступности данных и общую консоль управления. Это позволит использовать TagmaStore USP как ядро для построения самых сложных и крупных центров обработки данных.

Физически внешние системы хранения данных могут подключаться к TagmaStore USP напрямую (через FC-порты), через SAN или через расширители (рис. 2). Подключение серверов и мэйнфреймов также осуществляется как напрямую (через Fibre Channel/FICON/ESCON порты), так и через SAN.

При подключении, вне зависимости от способа, внешние системы хранения данных представляются пользователю полностью интегрированными с внутренним массивом дисков USP в виде единого пула LUN. Данная функциональность обеспечивается ПО Hitachi Universal Volume Manager. Этот программный продукт вместе с технологией

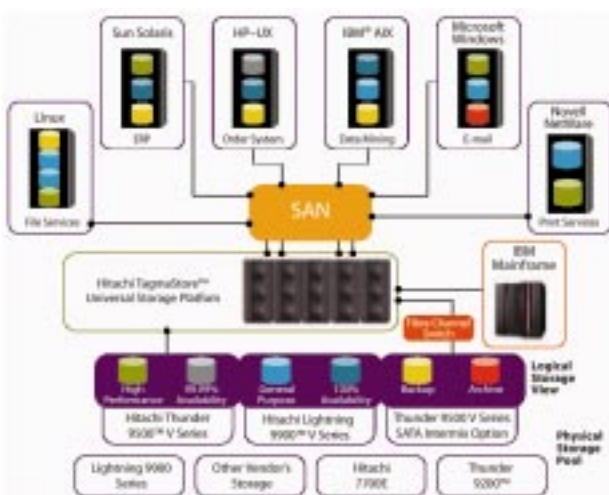


Рис. 1. Множество серверов и мэйнфреймов имеют доступ к консолидированным гетерогенным многоуровневым ресурсам хранения, управляемым TagmaStore USP.

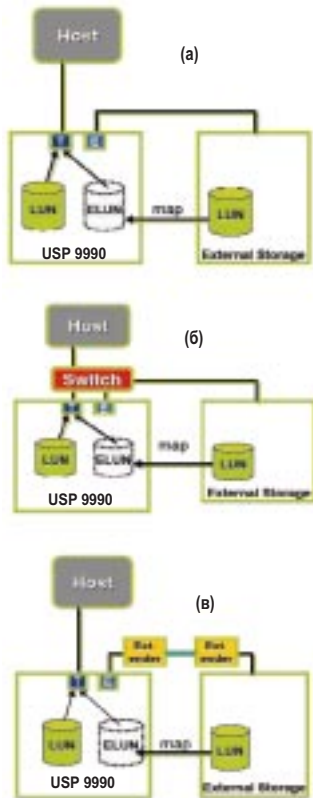


Рис. 2. Способы возможного подключения внешних систем хранения: напрямую (а); через FC-коммутатор (б); через расширитель (в).

Host Storage Domains (позволяющей разделять физические порты дискового массива на множество виртуальных портов и, кстати, доступной на всех моделях дисковых массивов Hitachi Data Systems) является основой мощного механизма виртуализации TagmaStore.

Объединение в один пул внешней и внутренней дисковой памяти с общей функциональностью и качеством обслуживания (QOS, Quality of Service) сокращает непроизводительные расходы на управление и предоставляет значительные возможности для повышения эффективности использования существующих устройств хранения.

Более того, возможности виртуализации TagmaStore USP позволяют существенно снизить расходы на программное обеспечение для управления данными. Например, в стандартной ситуации каждое устройство хранения имеет свое собственное ПО для тиражирования данных (моментальные снимки, репликация и др.). В случае USP мы можем использовать программные продукты, установленные на USP, для тиражирования данных, расположенных на внешних устройствах хранения. Тем самым устраняется необходимость применения нескольких программных продуктов с одинаковой функциональностью, и, соответственно, приобретения лицензий на их использование, обновления и заключения контрактов на сопровождение. Немаловажной с точки зрения сокращения эксплуатационных затрат является и возможность одновременного распространения последних усовершенствований ПО на все ресурсы хранения данных.

Все вышеперечисленное в результате обеспечивает снижение совокупной стоимости владения. По данным расчета ТСО для различных конфигураций TagmaStore (и, соответственно, различных объемов инвестиций), проведенного компанией ITCentrix, независимым поставщиком ПО и услуг, снижение совокупной стоимости владения для корпоративных пользователей составляет целых 40%. Так, для системы с емкостью 900 Тбайт и 2700 Тбайт внешних устройств хранения показатель возврата инвестиций составил 425% за 3 года.

“Это меняет саму динамику отрасли хранения данных”, – прокомментировал полученные результаты Дейв Велланте (Dave Vellante), главный исполнительный директор компании ITCentrix.

Важной функцией, также реализуемой ПО Universal Volume Manager вместе с ПО Volume Migration, является поддержка миграции данных в неоднородной среде хранения. Это позволяет управлять размещением данных на разных уровнях среды хранения на основе заданных правил, или, другими словами, управлять жизненным циклом данных. В частности, обеспечивается сохранение данных мэйнфреймов на экономичных системах хранения среднего класса, например на недорогих дисковых массивах, использующих диски SerialATA.

В будущем планируется поддержка USP корпоративных систем управления контентом (ECM, Enterprise Content Management). Таким образом, на базе USP будет проводиться агрегирование всех источников информации и протоколов в масштабах организации и дальнейшей оптимизации размещения контента (информационного содер-



Рис. 3. Через разделение на партии можно создать до 32 Private Virtual Storage Machines из внутренней и подключаемых систем хранения, каждая с выделенной емкостью, кэшем и портами.

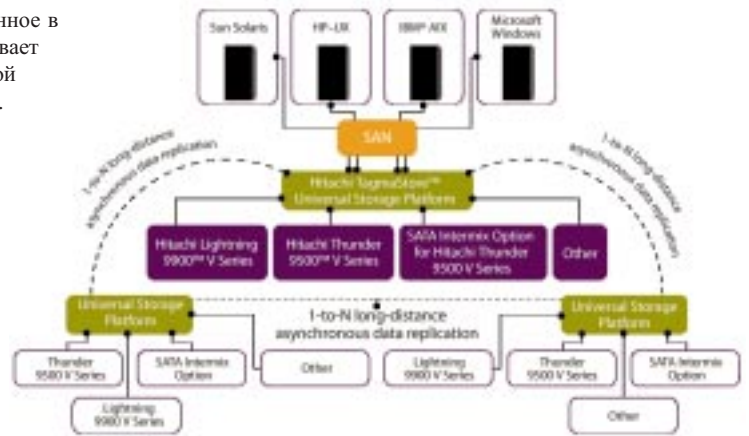


Рис. 4. Поддерживая на современном уровне реплицирование данных в географически удаленных центрах, USP 9990 удовлетворяет большинству требований восстановления после катастроф.

жимого) на разных уровнях хранения, включая и не-дисковое хранение, например, ленты.

Организация единого пула ресурсов хранения дает еще одну дополнительную возможность – выравнивание нагрузки. Используя ПО Hitachi Volume Migration (ранее известное как Hitachi CruiseControl), можно активно настраивать и балансировать рабочую нагрузку в пределах общей динамической области памяти. При этом можно балансировать нагрузку как внутренней памяти, так и внешней, подключенной в динамике и реальном времени, и это является прозрачным для конечных пользователей и приложений. Результат – повышение доступности данных, а также гарантия, что данные оптимально расположены для самого оперативного доступа.

Второе важное нововведение – возможность деления системы хранения данных на логические разделы. До настоящего момента эта технология применялась в основном на high-end серверах и мэйнфреймах, теперь она доступна и на системах хранения. ПО Virtual Partition Manager распределяет внешние и внутренние физические ресурсы хранения, включая порты, кэш и диски, выделяя независимо управляемые “виртуальные машины” (Private Virtual Storage Machine – PVSM). Максимальное количество PVSM – 32 (рис. 3). Разделы можно динамически изменять в соответствии с требованиями качества обслуживания (Quality of Service), исходя из меняющихся требований бизнеса и приоритетов приложений. В то же время, для хост-системы Private Virtual Storage Machine выглядит как собственная, отдельная система хранения данных. Таким образом обеспечивается децентрализованное использование и управление выделенными для хоста ресурсами.

Разделение на логические разделы позволяет в максимальной степени приблизиться к модели ресурсного управления. Сочетание Hitachi Storage Area Manager ПО и Virtual Partition Manager обеспечивает: динамическое выделение ресурсов “в нужное время в нужном месте”; полную детализированную отчетность по использованию ресурсов; максимально эффективное использование ресурсов хранения и др.

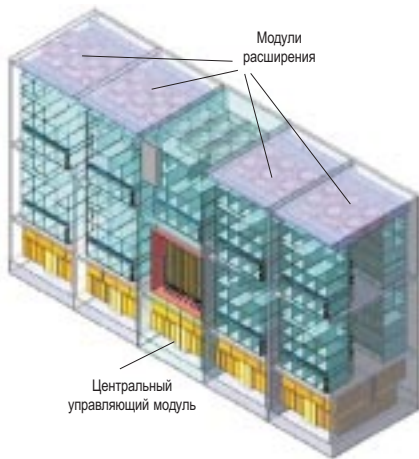


Рис. 5. Конструктивная архитектура TagmaStore USP.

Третье. HDS значительно расширила возможности по асинхронной репликации данных в гетерогенной среде (рис. 4). Эта функциональность обеспечивается ПО Universal Replicator, представляющим собой развитие ПО асинхронного удаленного тиражирования Hitachi TrueCopy™. Впервые заявлен целый ряд новых возможностей, включая репликацию в гетерогенной среде, ведение журналов дисков, защиту от отказа канала, копирование с опросом наличия изменений ("pull" copying). Также Universal Replicator поддерживает сценарии обеспечения непрерывности бизнеса для нескольких (больше двух) центров обработки данных, т.н. multi-data-center operation.

Таким образом, значительно расширяются возможности средств обеспечения непрерывности бизнес-процессов, которые теперь поддерживают любое расстояние, любые операции, а также внутренние и внешние системы хранения, что избавляет от необходимости устанавливать избыточные серверы или специализированные системы репликации данных.

Архитектура TagmaStore USP

Конструктивно TagmaStore USP, так же, как и Lightning 9980V, состоит из центрального управляющего модуля и 4 дополнительных модулей с дисками (рис. 5).

Архитектура TagmaStore USP представляет собой эволюционное развитие Lightning

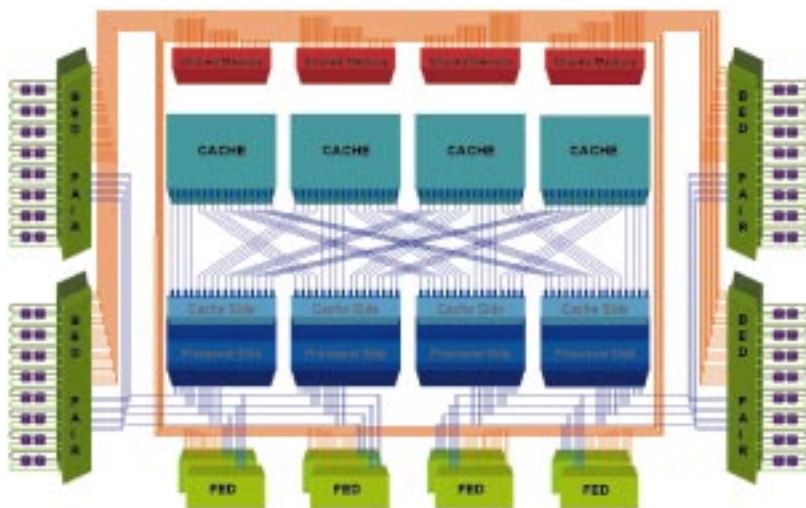


Рис. 6. Архитектура TagmaStore USP – это третье поколение Universal Star Network Crossbar Switch Architecture.

Табл. 1. Сравнение вариантов комплектации TagmaStore USP

Component	Description	USP100 Entry	USP600 Enhanced	USP1100 High-performance
Controller				
Basic Platform Unit				
	Integrated Control/Array Frame	1	1	1
	Number of Switches	2	4	4
Universal Star Network™ Crossbar Switch				
	Data Bandwidth (GB/sec)	17	34	68
	Control Bandwidth (GB/sec)	6.5	6.5	13
	Aggregate Bandwidth (GB/sec)	23.5	40.5	81
Data Cache				
	Cards	2	2	4
	Base Memory (GB)	4	16	32
	Maximum (GB)	64	64	128
Control Memory				
	Cards	2	2	4
	Base Memory (GB)	3	3	4
	Maximum (GB)	6	6	6
Front-end Directors Cards				
	1-2	1-6	1-6	
	Fibre Channel Ports	0-64	0-192	0-192
	Virtual Ports	до 8,192	до 24,576	до 24,576
	FICON Ports	0-32	0-48	0-48
	ESCON Ports	0-32	0-96	0-96
	NAS Blade™/Ports	0-2/0-8	0-4/0-16	0-4/0-16
Back-end Directors				
	Type	Standard	Standard	Standard
	Numbers	1	2	2,3,4
Logical Devices Supported				
		16,384	16,384	16,384
Array Frames				
Array Frames				
	Number	0,1	1,2	1,2,3,4
Hard Disks (GB)				
	Type	73, 146, 300	73, 146, 300	73, 146, 300
	Number	5-256	64-512	129-1152
Spare Drives per System				
	Minimum/Maximum	1/4	1/8	1/16
Internal Capacity (TB)				
	Minimum (73GB disks)	0.286	6.864	9.152
	Maximum (300GB disks)	74	148	330
Maximum Usable Capacity RAID-5 (TB)				
	Open Systems	62.4	128.8	287.8
	zOS-compatible	58.6	117.2	270.4
Maximum Usable Capacity RAID-1+ (TB)				
	Open Systems	36.4	72.8	165
	zOS-compatible	31.6	63.0	143.6
External Host Support				
External Capacity				
		32PB	32PB	32PB
Private Virtual Storage Machines				
		32	32	32
High Availability				
	Hi-Track® Support Package	LAN/Modem etc.	Standard	Standard
	Enhanced Support Package	2nd SVP	No	Optional

9900 и 9900V – третье поколение коммутируемой архитектуры Universal Star Network Crossbar Switch Architecture (рис. 6)

Ядром управляющего модуля является интеллектуальный коммутатор, который состоит из двух отдельных сетей: cache hierarchical star network и control memory hierarchical star network. Первая – сеть передачи данных сети к/от глобального кэша данных. Вторая – отвечает за обмен управляющей информацией между процессорами и управляющей памятью (control memory). Таким образом, поток управляющей информации полностью отделен от потока данных.

Cache hierarchical star network строится на основе кэш-коммутаторов, которых в TagmaStore USP может быть до 4. Каждый кэш-коммутатор имеет 16 путей к модулям кэш-памяти 16 путей в сторону процессоров (в 2 раза больше, чем Lightning 9900 V) и представляет собой специально разрабо-

танный перекрестный коммутатор, который функционирует как комбинация мультиплексора, арбитра пути и неблокируемого сетевого коммутатора. Всего 64 пути в процессорную сторону и 64 пути к модулям кэша поддерживают одновременное выполнение до 128 параллельных операций кэш-памяти.

Control memory hierarchical star network имеет более простую структуру – это комбинация соединений точка-точка между модулями управляющей памяти и front-end / back-end директорами. Управляющая память содержит информацию о статусе, местоположении и конфигурации кэша, данных в кэше и конфигурации системы в целом (так же, как и другая информация, связанная с операционным состоянием системы). Области управляющей памяти попарно зеркалированы. Front-end директора (FED) – это контроллеры, отвечающие за внешние интерфейсы и определяющие тип и количество соединений с серверами/SAN/мэйнфреймами. Back-end директора (BED) – это контроллеры, отвечающие за интерфейсы к дискам и определяющие количество FC-AL петель, идущих к пулу внутренних дисков.

Cache hierarchical star network обеспечивает передачу данных со скоростью 68 Гбайт/с, control memory hierarchical star network – со скоростью 13 Гбайт/с. Таким образом, агрегированная производительность коммутатора – 81 Гбайт/с.

Рассмотрим более подробно front-end и back-end директора, соответственно определяющие связь USP с серверным пулом и внешними устройствами хранения данных, а также основные характеристики USP как системы хранения.

TagmaStore USP поддерживает до 6 front-end директоров, которые обеспечивают следующие внешние интерфейсы: Fibre Chan-

nel, ESCON, FICON, Ethernet. В будущем планируется поддержка iSCSI.

Общее число портов – до 192 Fibre Channel, до 48 FICON-портов, до 96-ESCON-портов. Также USP поддерживает до 32 1Gbps Ethernet портов для NAS-устройств с CIFS и NFS протоколами.

USP может иметь от 1 до 4 back-end директоров. Каждый back-end директор поддерживает 16 2Gbps Fibre Channel Arbitrated Loop (FC-AL) петель, или суммарно – 64 петли. Как заявляет разработчик, архитектура back-end директоров полностью изменена в сравнении с предыдущей линейкой (модели 9980V/9970V), что позволило увеличить пропускную способность FC-петель, а также обеспечить поддержку back-end директорами более высоких скоростей в будущем в сравнении с существующей 2Gbps.

Таким образом, при максимально возможном числе дисков – 1152 на каждую петлю (из 64) будет приходиться 18 накопителей. Из этих показателей можно легко оценить максимальную потоковую и случайную производительность ввода/вывода со стороны back-end директоров или непосредственно массива дисков. В ближайшее время должны появиться результаты измерений на тестах SPC.

USP строится на дисках с емкостью/групп 72 Гбайт/15К и 146 Гбайт/10К. В ближайшее время должны появиться диски 300 Гбайт/10К.

Комплектация и конструктивная реализация

Конструктивно USP предлагается в трех вариантах исполнения (табл. 1):

- Entry (USP100) – управляющий модуль и модуль расширения, опционно;
- Enhanced (USP600) – управляющий модуль и 1-2 модуля расширения;
- High-performance (USP1100) – управляющий модуль и 4 модуля расширения.

Центральный управляющий модуль содержит:

- до 128 64-bit симметричных микропроцессоров;

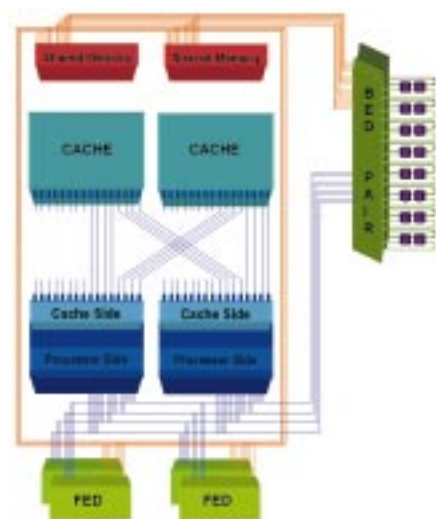


Рис. 7. Внутренняя архитектура начальной модели TagmaStore USP.

- неблокирующий коммутатор 3-го поколения;
- до 128 Гбайт cache memory (256 Гбайт – в будущем);
- до 6 Гбайт control memory (12 Гбайт – в будущем);
- 16 адаптерных слотов (4 – резерв для front-end директоров);
- 8 дисковых полок (каждая по 16 дисков);
- сервисные компоненты.

USP 100 (рис. 7) эквивалентен по производительности модели 9980V. USP 600 (рис. 8) соответствует удвоенной производительности модели 9980V. High-performance конфигурация USP 1100 соответствует примерно учетверенной производительности модели 9980V (см. рис. 6).

Доступность

TagmaStore™ USP поставляется HDS как напрямую, так и через OEM-партнеров – Sun Microsystems и HP. Не лишним будет упомянуть, что при OEM поставках система будет иметь другое название: Sun StorEdge 9990 и XP12000 соответственно. Отметим также, что доступность всей заявленной функциональности, как и список поддерживаемых платформ и систем хранения, зависит от того, “под каким именем” будет поставаться система, и требует уточнения в каждом конкретном случае. Аппаратная часть системы доступна к поставке уже с сентября этого года от всех трех компаний.

Надо полагать, что до 2 кв. 2005 г. все вопросы, связанные с поддержкой USP подключения внешних систем хранения от других производителей, будут урегулированы.

На текущий момент о взаимной сертификации TagmaStore и “старых” Shark’ов заявили IBM и HDS. До конца года ожидается сертификация подключения T3 как внешнего устройства.

В соответствии с roadmap HDS, этапность введения функциональности USP TagmaStore представлена на рис. 9.

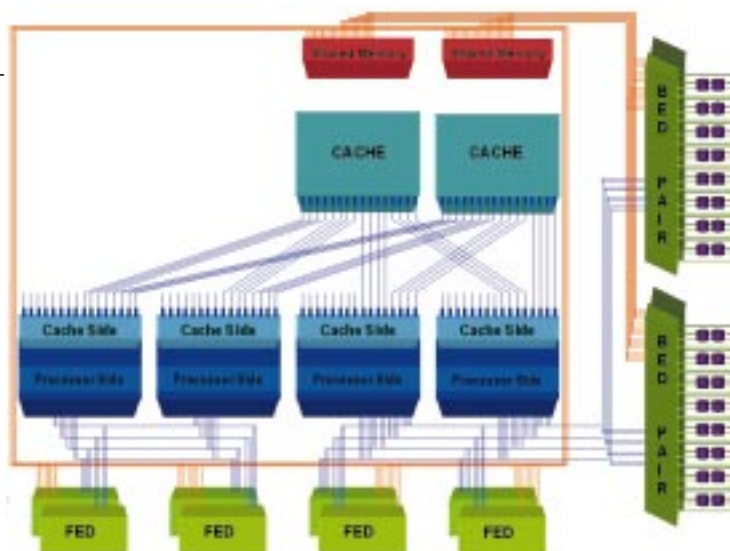


Рис. 8. Внутренняя архитектура расширенной модели TagmaStore USP.

Заключение

Компания HDS сделала очень важную заявку на рынке для упрочения своего положения. Возможно даже, что предложение TagmaStore USP сыграет определяющую роль для дальнейшего развития индустрии систем хранения данных. Осталось дождаться ответных шагов со стороны других производителей систем хранения и реакции самого рынка.

Наталья Гребеникова,
эксперт компании “Инфосистемы Джет”

Компания “Инфосистемы Джет” – ведущий российский системный интегратор, поставщик основных компонентов информационной инфраструктуры для крупных организаций и предприятий.

Ключевые области специализации компании – проектирование, внедрение и обслуживание высокопроизводительных серверных комплексов, систем и сетей хранения данных (SAN).

Москва, ул. Б. Новодмитровская, 14/1
тел.: (095) 411-7601
факс: (095) 411-7602
storage@jet.msk.su
www.jet.msk.su

	September 7th 2004	End December 2004	During 2005
Universal Storage Platform	Web RAW		Web RAW
	T3 & M3000 Data		M3000 HDDs – February 05
	WORM Addressing	FullRA MP solution	High performance Carbon
	SPS External Capacity	T30 - 3rd party Ext. Storage	30 FICON & 6000 FICON QDR
	1000B Cache		1000B Shared Memory-QDR
	1100 Drives (WE)		2000B Cache – QDR
	600 Shared Memory		QCSI (not in use)
	FC, F-FC, FICON, FICON		No Performance Enhancements
	External Hi-cap Storage		Recovery of 2004 – 2005
	Competitive Big Migration		Copy-on-Write Snapshot
TagmaSoft	TrueCopy	TrueCopy To Ext. Values	
	SkidStorage		
	Cache Resiliency (P/A)	HSBC	
	FCcopy		
	Data Retention Utility (DRU)	Data Retention Utility Enhancement	
	Storage Navigator	Volume Migrator Enhanced	
	Performance Monitor		
	Volume Migrator (QDR)		
	Priority Access		
	Resource Manager	Flash Copy W	
HCCommand Modules	Universal Replicator w/DC		
HCLM			
HCLM	Virtual Partition Manager		

Рис. 9. Этапность введения функциональности USP TagmaStore.