

Nutanix для “тяжелых” приложений

В 2016 г. компания Nutanix¹⁾ представила два новых релиза своей платформы, значительно расширив ее возможности, а в начале ноября 2016 г. объявила о сертификации SAP® Business Suite powered by SAP NetWeaver® on Linux для выполнения на платформе Nutanix Enterprise Cloud Platform с AHV-виртуализацией. Одновременно было анонсировано, что пользователи, выполняющие SAP Business Suite, будут иметь возможность использовать Nutanix Acropolis Block Services для организации внешнего хранилища при работе с реляционными БД, включающими: SAP Adaptive Server® Enterprise (SAP ASE), Oracle DB, Microsoft SQL Server и IBM DB/2 на существующих bare-metal или виртуализованных серверах. Развитие этого направления, возможно, сделает 2017 год для Nutanix одним из самых значимых по числу поддерживаемых “тяжелых” приложений, включая in-memory обработку БД.



Роман Хмелевский — старший системный инженер, Nutanix Inc.

Предыстория

“Конвергентные системы” (или интегрированные системы, *прим. ред.*) как термин появились на рынке несколько лет назад²⁾ и были представлены решениями Oracle, VCE (Cisco+EMC) vBlock, NetApp/Cisco FlexPod и рядом других, но представляли собой, по сути, классические архитектуры, состоящие из серверов, систем хранения и SAN (или на базе других протоколов, отличных от FC, *прим. ред.*), просто собранных для удобства покупки в единый rack и продававшихся как уже соединенное и настроенное единое целое. Основное их преимущество — быстрота установки и развертывания приложений. Такую систему сейчас принято называть “конвергентные 1.0”. Они, безусловно, сыграли свою роль, но не привели к прорыву, так как наследовали все присущие традиционным системам недостатки архитектуры.

“Гиперконвергентными”, “конвергентными 2.0” или “web-scale” принято называть системы, где “конвергентность” (слияние, объединение) затрагивает уже архитектурную часть решения.

Одним из толчков для создания подобных систем была необходимость обеспечивать масштабирование решения платформы

для приложений, при продолжающемся росте ИТ-инфраструктуры, более плавно и с меньшим числом проблем, чем сегодня.

Несколько лет казалось, что проблема масштабирования решена с помощью массового перехода и переноса приложений в виртуализованную среду, “отвязки” их от физической платформы сервера и облегченной миграции по хост-серверам датацентра on-demand. Однако практика показала, что это лишь частично ослабило остроту проблемы, но не решило ее в целом. Основная проблема, вышедшая на первый план после перехода в виртуализованную среду и к cloud-модели, это — недостаток гибкости и линейности масштабирования решения.

К тому же, производители оборудования серверов, систем хранения и сети передачи данных идут по пути “новых мэйнфреймов”, привязывая к себе и своим “фичам” пользователей, делая для них, невозможным (или очень дорогим) отказаться от используемых платформенных решений. Переход же в Open Source-лагерь часто невозможен из-за недостатка компетенции на стороне ИТ-персонала компании и проблем с поддержкой.

Таким образом, “проблема Scale-Out”, то есть масштабирования решения, сегодня стоит почти так же остро, как и несколько лет назад, до массового входа в отрасль систем виртуализации и “облачности”.

Становится очевидным, что дальнейший прогресс тут сдерживается “на архитектурном уровне” и связан с тем, что “классическая” архитектура (то есть архитектура, содержащая отдельно серверы, отдельно — системы хранения, централизованно хранящие данные VM и приложений, работающих на серверах, и отдельно — сеть передачи данных — SAN, соединяющая их) сами по себе плохо приспособлены для работы в виртуальной среде и требующегося сегодня уровня масштабирования ИТ-решений.

Nutanix решает проблему scale-out за счет того, что каждый узел “кластера” Nutanix — это индивидуальный *compute-and-storage* блок, оснащенный всем необходимым для своей работы, CPU, памятью и диском. Это своеобразный “кубик LEGO” датацентра будущего.

При необходимости увеличить мощность такого датацентра пользователь просто добавляет в кластер нужное число нужного ему типа “кубиков”-нодов Nutanix, и поскольку каждый из них приносит в общий “ресурсный пул” датацентра и CPU, и память, и емкость хранения, вы можете не беспокоиться, что добавление серверов

приведет к тому, что не будет хватать пространства хранения или увеличение объема и пропускной способности системы хранения упрется в насыщение канала SAN или иной сети передачи данных, и так далее.

Такая независимость и простота соединения “кубиков” нашего “конструктора датацентров” дает нам гибкость построения, и гарантию, что мы никогда не будем вынуждены полностью перестраивать инфраструктуру датацентра, выбрав для него такую архитектуру.

Одними из ключевых преимуществ платформы Nutanix являются:

- гипермасштабируемость — от нескольких единиц до сотен тысяч узлов без увеличения сложности администрирования;
- снижение капитальных и эксплуатационных затрат на администрирование инфраструктуры;
- поддержка широкого класса корпоративных приложений и нагрузок;
- более простое: достижение отказоустойчивости, устранение узких мест производительности, реконфигурирование в случае отказа/сбоя.

Архитектура

Программное обеспечение

Платформа Nutanix Xtreme Computing Platform (XCP) является программно-определяемой, основу которой составляют два программных продукта: Nutanix Acropolis и Nutanix Prism.

Nutanix Acropolis — горизонтально-масштабируемая фабрика данных для их хранения, обработки и виртуализации.

Acropolis сочетает в себе атрибуты SDS-хранения (software-defined storage, SDS) со встроенной виртуализацией для создания “под ключ” решения гиперконвергентной инфраструктуры. Nutanix Acropolis состоит из трех основных компонентов:

- *распределенная фабрика хранения (Distributed Storage Fabric, DSF)*. Эта компонента обеспечивает основные сервисы хранения данных: распределенную файловую систему, дедубликацию данных, автоматизированное многоуровневое хранение и моментальные снимки. Она поддерживает множество протоколов доступа гипервизора к хранимым данным, таких как NFS, SMB 3.0 или iSCSI, что позволяет объединять большое количество разрозненных ра-

1) Компания Nutanix, образованная в 2009 году в Сан-Хосе, Калифорния, в «Кремниевой Долине», занимается разработкой и производством единственного продукта — гиперконвергентной платформы Nutanix Virtualized Platform, и на сегодня работает по всему миру, насчитывает более 2500 сотрудников, около \$1 млрд в год объема продаж, и имеет в своем активе свыше 3800 клиентов, в том числе 150 аккаунтов с продажами на миллион долларов и выше.

По оценке IDC, компания в настоящее время занимает около 50% рынка гиперконвергентных (web-scale) систем. Компания вышла на IPO на бирже NASDAQ в сентябре 2016 года и ее рыночная капитализация, по текущей оценке, составляет около \$3,5 млрд.

2) При подготовке публикации использовались перевод и оригинал книги Nutanix Bible (<http://nutanix.ru/>, <http://nutanixbible.com/>), создаваемой и поддерживаемой Стивом Пойтрасом, архитектором и евангелистом компании Nutanix.

бочих нагрузок на единой платформе виртуализации;

- фабрика мобильности приложений (*App Mobility Fabric*). Эта компонента обеспечивает размещение ВМ, миграцию ВМ и преобразование виртуальных машин, а также кросс-гипервизорную высокую доступность и интегрированную катастрофоустойчивость;
- гипервизор *Acropolis (Acropolis Hypervisor, AHV)*. В то время как DSF полностью поддерживает традиционные гипервизоры, такие как VMware vSphere/ESXi и Microsoft Hyper-V, Acropolis также включает в себя и собственный свободный гипервизор, основанный на коде хорошо известного гипервизора Linux KVM, что обеспечивает повышенную безопасность и самовосстановление, основанные на SaltStack и на управлении виртуальными машинами корпоративного класса. При этом поддерживаются все возможности, которые предлагает гипервизор (HA, FT, Live Migration, Distributed Switch, DRS), а с точки зрения “администратора виртуальной инфраструктуры” – все остается в привычном ему интерфейсе управления (SCVMM или vCenter). Для собственного гипервизора (Acropolis Hypervisor) Nutanix разработал интегрированный веб-интерфейс управления, встроенный в Nutanix Prism и построенный на технологии HTML5.

Nutanix Prism является собственной разработкой Nutanix и дает администраторам простой способ управлять своими виртуальными средами за счет объединения нескольких аспектов управления центрами обработки данных в единое решение потребительского класса. Nutanix Prism использует технологию машинного обучения для извлечения больших объемов информации о системе и генерирования эффективных механизмов оптимизации всех аспектов управления виртуальной инфраструктурой.

Аппаратное обеспечение

С аппаратной точки зрения, узел кластера Nutanix – это сервер с локально подключенными в него жесткими дисками (SSD и SATA), доступ к которым для гипервизора виртуализации на нем осуществляется через специальную виртуальную машину (CVM – Controller VM), существующую под тем же bare-metal-гипервизором (например, VMware ESXi, MS Hyper-V или Acropolis Hypervisor, AHV). В CVM работают сервисы, обеспечивающие размещение поверх физических жестких дисков слоя кластерной распределенной файловой системы, объединяющей все диски кластера в общее пространство хранения и обеспечивающие отказоустойчивость для хранимых данных. На этом слое кластерной файловой системы располагается datastore гипервизора, а внутренние средства Nutanix обеспечивают отказоустойчивое избыточное распределение по нему данных и защиту от отказов оборудования без использования технологии RAID.

Nutanix использует стандартную платформу x86-сервера “стандартной архитектуры” (в настоящее время это серверная платформа производства Supermicro), не внося в нее каких-то “брендируемых” изменений, не навязывает собственные компоненты, например, жесткие диски

“от вендора” или аппаратные “акселераторы” чего-нибудь, и является полностью “софтверным решением” (Software defined). Тем не менее, для enterprise-рынка и оказания полноценной поддержки и гарантии мы вынуждены поставлять наше программное решение, устанавливаемое поверх валидированной платформы, и в настоящее время в продакшне не поддерживается установка Nutanix на стороннее “железо”, например, на уже существующее у потенциального клиента серверное оборудование сходного типа.

Однако на рынке наше решение предоставляют и другие вендоры, работая с нами как OEM-партнеры, например – Dell, выпустивший свою новую линейку конвергентных систем Dell XC, которые представляют собой ПО Nutanix, стоящее на сервере Dell PowerEdge и продаваемое через дистрибуторский канал самого Dell, а также Lenovo, в 2016 году представившая аналогичный OEM-продукт Lenovo NX на базе серверной платформы Lenovo и ПО Nutanix.

Наши “кубики LEGO” неодинаковы. За счет этого достигается оптимальность подстройки под то, что требуется из них построить. В составе линейки Nutanix есть узлы-ноды общего применения (серия 3000); есть capacity nodes (серия 6000) – узлы, имеющие больше емкости хранения и меньше вычислительных “мускулов” CPU и RAM; есть узлы для специализированной VDI-нагрузки, например, для тяжелого CAD/CAM, 3D-моделинга, видеообработки (модели 3175 и 3055G), для мощных баз данных и систем Exchange с тысячами ящиков (серия 8000). Большинство систем может быть сконфигурировано как All-Flash, включая использование SSD вплоть до 3,8TB объемом каждый.

Важным при этом является то, что все эти узлы могут составлять “гетерогенный кластер”, можно соединять в кластер любой набор узлов, необязательно одинаковых. Если в инфраструктуре есть задача преимущественно хранения – добавьте узлы 6000-й серии; требуется платформа для мощной СУБД – плюс еще один или два узла 8000. Нужно запустить VDI для пользователей с тяжелыми CAD или производительной HD графикой – модели NX-3175/3055. При этом для приложений и виртуальных машин остается доступным все дисковое пространство всех дисков всех входящих в кластер отдельных его узлов, даже разнородных.

Решение Nutanix – это конвергентное решение, объединяющее хранилище и вычислительные модули, которое использует локальные компоненты и создает распределенную платформу виртуализации (или virtual computing platform). Решение представляет собой поставляемое в едином

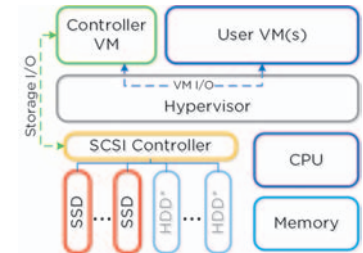


Fig. 1. Логическая организация типового узла.

комплекте аппаратное обеспечение и ПО. Каждое устройство (Block) состоит из 2 узлов (Nodes) (для устройств 6000/7000 серии) или 4 узлов (для узлов 1000/2000/3000 серий) и имеет размер 2U.

Также существуют компактные модели в формате 1U и высокопроизводительные узлы 8000-й серии, с 24 локальными дисками в вычислительном узле-ноде высоты 2U.

Каждый узел кластера Nutanix обеспечивает запуск стандартных гипервизоров (на данный момент это ESXi, AHV, Hyper-V), а также Nutanix Controller VM (CVM). Именно на CVM выполняется ПО Nutanix и обслуживаются все операции I/O для гипервизоров и ВМ, размещенных на узле. Для узлов Nutanix с запущенной VMware vSphere, контроллер SAS, который управляет устройствами SSD и HDD, напрямую презентуется в CVM, используя VM-Direct Path (Intel VT-d). В случае с Hyper-V устройства хранения напрямую презентуются в CVM. На рис. 1 дан пример логической организации типового узла.

Вместе, группа узлов Nutanix формирует распределенную платформу, называемую Acropolis Distributed Storage Fabric (DSF) (панель, Nutanix Distributed Filesystem, NDFS). DSF предоставляется гипервизору как централизованное хранилище данных, однако все операции ввода/вывода (I/O) выполняются локально для обеспечения максимальной производительности. Пример того, как узлы Nutanix формируют DSF/NDFS – на рис. 2.

Основные применения

Основные области применения, где играет Nutanix, это, например: системы серверной виртуализации (private/hybrid cloud), VDI-инфраструктуры, платформы для обработки Big Data (Hadoop, Splunk и др.), Business Intelligence, ERP, включая SAP и Oracle, СУБД Oracle и MS SQL Server, систем Unified Communications (MS Exchange/Lync/Avaya), распределенных и отказоустойчивых инфраструктур и решений Disaster Recovery.

“Не играет” Nutanix там, где неприменима виртуализация, где приложения требуют

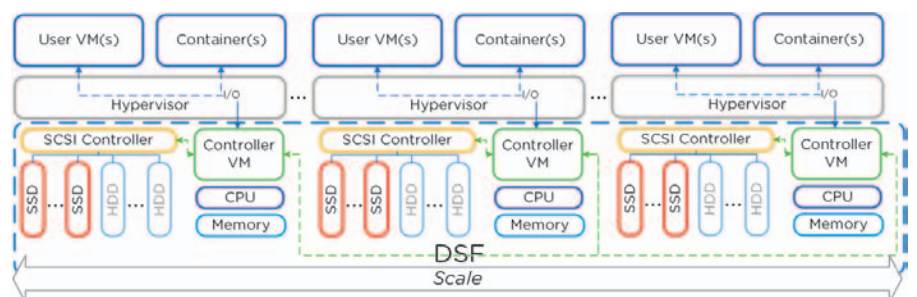


Fig. 2. Пример, как узлы Nutanix формируют NDFS.

не-x86 платформы (SPARC, IBM Power, etc.) и где используется преимущественно “холодное” хранение (архивы, малоиспользуемые данные резервных копий).

Также Nutanix не может “сделать невозможное”, и “собрать” из нескольких под-серверов один “суперсервер”, объединив ядра CPU и память отдельных нод в один огромный “мегасервер” для сверхмощной одиночной задачи. Дисковая емкость всех индивидуальных серверов, однако, объединяется.

Внутреннее ПО Nutanix обеспечивает работу распределенной кластерной системы по всем узлам кластера, создание и мониторинг хранилища данных (datastore container) в гипервизор, мониторинг и управление, контроль за производительностью, автосбалансирование в случае отказов, ребалансировку ввода-вывода, а также ряд сервисов, обычно реализуемых внешними службами, такие, как создание резервных копий, обеспечение репликации данных между кластерами (синхронной и асинхронной), а также организацию географически распределенного синхронного кластера (metro-available cluster).

Таким образом, используя платформу Nutanix для применяемого в вашей компании гипервизора виртуализации, можно построить компактную, гибкую, экономную, с точки зрения эксплуатационных расходов, высокопроизводительную, эффективную платформу, не имеющую ограничений при необходимости роста, построенную в значительной мере на открытом коде и аппаратных решениях, и позволяющую компании и ее IT-инфраструктуре расти вслед за требованиями бизнеса.

Оценка эффективности

В августе 2015 г. аналитическая компания IDC опубликовала результаты обследования 13 организаций, развернувших решения Nutanix (*Quantifying the Business Value of Nutanix Solutions*, Authors: Matthew Marden, Eric Sheppard, August 2015, IDC). Цель исследования — количественно оценить эффективность решений Nutanix в сравнении с предшествующими им IT-инфраструктурами.

IDC проинтервьюировало 13 организаций для данного исследования: из них 10 организаций использовали решения Nutanix XCP и в 3-х организациях были развернуты Dell XC Web-scale устройства (совместно именуемые «решения Nutanix»). Опрос проводился в апреле и мае 2015 года.

В опросе участвовали IT-менеджеры и бизнес-лица, принимающие решения в этих организациях. Эти организации либо мигрировали свои нагрузки с традиционной трехуровневой (серверы, систем хранения, сети) инфраструктуры центров обработки данных, либо развернули новые рабочие нагрузки на решениях Nutanix. Опрошенные организации варьировались от поставщиков услуг — сервис-про-

Табл. 1. Демография интервьюируемых организаций

	Average	Median	Range
Число работающих	4094	1500	45–18 000
Число IT-сотрудников	210	62	2–1500
Число внутр. IT-польз.	3,738	1,260	45–18 000
Число бизнес-прилож.	185	40	12–1500
Объемы данных (Тбайт, гав)	857	200	20–5120
Страны	США, Канада, Великобритания, Франция, Швеция, Швейцария, Австрия		
Отрасли	Производство, страхование, фин. услуги, здравоохранение, организация автогонок, правительственные орг-ции, игры, розничная торговля, поставщики услуг		



Рис. 4. Средняя ежегодная экономия на одну организацию составила \$2,22 млн (ист. IDC, 2015).

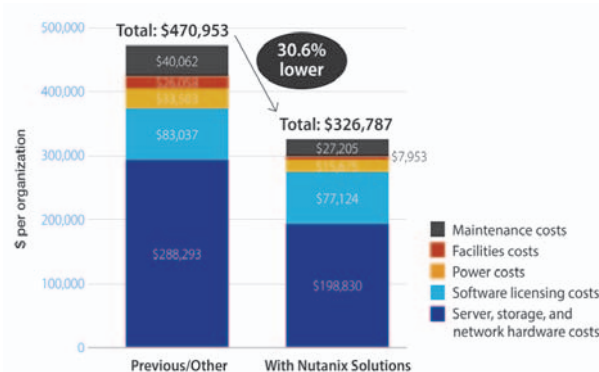


Рис. 5. Средние ежегодные затраты на IT-инфраструктуру уменьшились на 30,6% (ист. IDC, 2015).



Рис. 6. Среднее улучшение основных показателей при развертывании/управлении Nutanix IT-инфраструктуры (ист. IDC, 2015).

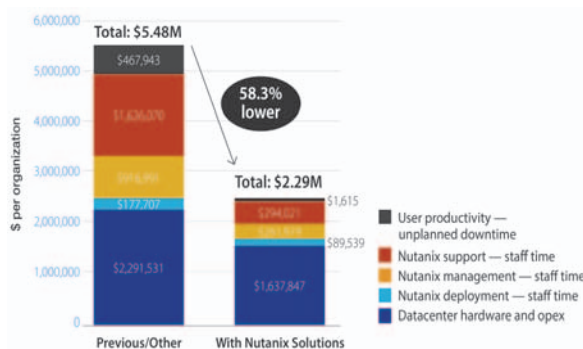


Рис. 7. Среднее уменьшение ТСО на организацию за пятилетний период при развертывании Nutanix IT-инфраструктуры — более 58% (ист. IDC, 2015).

вайдеров — с штатом менее чем 100 сотрудников до более крупных организаций — до 18 000 сотрудников (в среднем 4094 человек). Количество развернутых бизнес-приложений варьировалось в диапазоне от 12 до 1500 (в среднем по 185) — табл. 1.

Результаты исследования в обобщенном виде представлены на рис. 3–6. Средняя ежегодная экономия на одну организацию составила \$2,22 млн (см. рис. 3). Средние ежегодные затраты на IT-инфраструктуру уменьшились на 30,6% (см. рис. 4). Среднее улучшение основных показателей при развертывании/управлении Nutanix IT-инфраструктуры в сравнении с предыдущей — рис. 5. Среднее уменьшение ТСО на организацию за пятилетний период — более 58% (рис. 6).

Направления развития платформы Nutanix XCP

Еще одно направление, которое Nutanix развивает, это аналитика больших данных на базе hadoop-кластеров. Наш первый клиент в России — «Связной», и это как раз — поддержка hadoop. В «Связном» около 30 нод в hadoop-кластере, на котором обсчитывают разнообразие неструктурированные и структурированные данные, накопленные всей группой «Связной» (ритейлом, банковской группой и так далее).

Несколько причин, по которым виртуализация hadoop-узлов становится целесообразна:

— *увеличение производительности.*

По проведенным оценкам (*A Benchmarking Case Study of Virtualized Hadoop Performance on VMware vSphere 5*, VMware. <http://www.vmware.com/files/pdf/VMW-Hadoop-Performance-vSphere5.pdf>), виртуализованные узлы hadoop-кластера показывают большую производительность, чем при развертывании на bare-metal серверах;

— *большая утилизация ЦПУ.* Развертывание hadoop на “голом железе” позволяет достичь лишь 10–20%-ной утилизации ЦПУ. Виртуализация повышает гибкость и загрузку аппаратных средств;

— *большая эластичность и масштабируемость MapReduce.* Возможность динамического добавления и удаления узлов Hadoop в соответствии с текущей нагрузкой позволяют масштабировать конфигурацию на основании текущих потребностей и без излишней избыточности;

— *возможность сосуществования различных нагрузок,* например, выполнение пакетных заданий MapReduce можно запланировать в нерабочее время — ночью или в выходные дни, не снижая производительности, например, VDI-нагрузок;

— *возможность поддержки сервис-ориентированного управления* на основе приоритизации ресурсов и резервирования;

— *снижение зависимости от сбоев/отказов, нехватки ресурсов при реализации сложных проектов.*

Nutanix является партнером Hortonworks — крупнейшего «коммерческого» поставщика решений на Hadoop (<http://hortonworks.com/partner/nutanix/>).

В целом, Big Data — это один из самых быстро растущих секторов нашего бизнеса, и мы активно это направление развиваем и пропагандируем. Наша архитектура (много вычислительных узлов-нод с локально расположенными дисками для хранения данных) очень хорошо соответствует тому, что хочет «видеть» Big Data в качестве своей платформы.

Основные усилия Nutanix в настоящее время сосредоточены на расширении спектра поддерживаемых приложений, прежде всего, высоконагруженных корпоративных. В подтверждение этого в ноябре 2016 г. состоялся анонс сертификации SAP® Business Suite powered by SAP NetWeaver® on Linux для выполнения на

платформе Nutanix Enterprise Cloud Platform с AHV-виртуализацией. Одновременно было анонсировано, что пользователи, выполняющие SAP Business Suite, будут иметь возможность использовать Nutanix Acropolis Block Services для организации внешнего хранилища при работе с реляционными БД, включающими: SAP Adaptive Server® Enterprise (SAP ASE), Oracle DB, Microsoft SQL Server и IBM DB/2 на существующих bare-metal или виртуализованных серверах.

В 2017 г. ожидается расширение модельной линейки, направленное на поддержку новых производительных приложений enterprise-уровня. Это позволит разворачивать высокопроизводительные приложения реального времени с поддержкой in-memory технологий обработки БД. Также в обеспечение технологии in-memory обработки в новом релизе (1 кв. 2017 г.) ожидается поддержка технологии vNUMA.

В начале 2017 г. появится новое поколение SSD на базе технологии Intel 3D Xpoint, обещающее тысячекратное снижение задержек по сравнению с сегодняшними флэш-накопителями и сравнимое со скоростями DRAM, а в конце 2017 г. — новое поколение DIMM более емких и меньшей стоимости. Использование их в конвергентных системах, подобных Nutanix, в которых данные хранятся не в SAN, а, фактически, рядом с CPU, с доступом к ним с минимальными задержками, станет еще одним значительным плюсом новых архитектур. Сегодня архитектура Nutanix и подобных ему гиперконвергентных web-scale систем готова к использованию этих новых прорывных продуктов, которые будут формировать IT-будущее в ближайшие годы.

*Роман Хмелевский,
Nutanix Inc.*



NAVICA DATA

Крупнейший поставщик NUTANIX
в России и странах СНГ

Авторизованный сервис
и техническая поддержка Nutanix
Учебный центр, лаборатория, PoC

ДИСТРИБЬЮТОР ИТ-ОБОРУДОВАНИЯ

ARGO
TECHNOLOGIE SA

СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ

NUTANIX
Your Enterprise Cloud Platform

ГИПЕРКОНВЕРГЕНТНЫЕ
ПЛАТФОРМЫ ВИРТУАЛИЗАЦИИ

ARISTA

ВИСОКОПРОИЗВОДИТЕЛЬНЫЕ
КОММУТАТОРЫ

GeoVision Ltd.
The Vision of Security

СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ

KEMP

БАЛАНСИРОВЩИКИ НАГРУЗКИ СЕРВЕРОВ

ООО «Навика Дата», www.navicadata.ru, sales@navicadata.ru, тел.: +7 495 150 39 93
109004, г. Москва, ул. Земляной Вал, 50А/8 с2