PLATFORMS

Семейство T-Blade V-Class. Обзор вычислительных модулей V205S и V205F

Семейство T-Blade V-Class. Обзор вычислительных модулей V205S и V205F

Ревизия: 2.0 от 01/05/2012

Оглавление

Семейство T-Blade V-Class

Обзор вычислительных модулей V205S и V205F

1. Вычислительные модули V205S и V205F

Вычислительные модули V-Class V205 разработаны компанией Т-Платформы, ведущим российским производителем HPC-решений, на основе новейших процессоров AMD Opteron™ 6200 «Interlagos» и ускорителей NVIDIA® Tesla™ серии M.

Модули V205 позволяют создавать разнообразные конфигурации вычислительных систем на базе шасси V5000 для научных и коммерческих приложений. Модули поставляются в двух версиях: стандартной толщины (V205S) и двойной толщины (V205F), и доступны в типовых конфигурациях и под заказ.

1.1 Позиционирование вычислительных модулей V205

Высокая вычислительная плотность

Полностью заполненное шасси V5000 в исполнении 5U в максимальной конфигурации содержит десять модулей V205S с 320 ядрами AMD Opteron™ (2560 ядер в стойке 42U), что позволяет создавать плотные высокопроизводительные системы при достаточно низкой стоимости. К примеру, высокоуровневая система TB2-XN в шасси 7U с 32 узлами на базе процессоров Intel® Xeon® 5500/5600, созданная компанией Т-Платформы в 2009 году, имеет 384 ядра и стоит более чем в два раза дороже. При этом узлы V205 могут оснащаться разнообразными процессорами: от четырехядерных моделей с фиксированной частотой 3,3 ГГц до 16-ядерных моделей с базовой частотой 2,3 ГГц.

Великолепная производительность в расчете на ватт потребляемой энергии

В сравнении с процессорами AMD серии 6100 «Magny Cours», новейшие процессоры серии 6200 «Interlagos» обеспечивают повышение производительности многопоточных приложений на 10-35% с аналогичным тепловыделением TDP в 115 Вт. Новые технологии энергосбережения, такие как TDP Power Capping, обеспечивают точечный контроль тепловыделения, позволяя размещать большее количество вычислительных узлов в средах с ограничениями по энергоснабжению и охлаждению комплексов.

Ускорение приложений за счет адаптеров GPU

Узел V205F обеспечивает поддержку одного ускорителя NVIDIA® Tesla™ серии M, что позволяет установить до 5 ускорителей в составе одного шасси V5000. Модель гетерогенных вычислений входит в стадию признания рынком, и многие производители ПО и академические заведения оптимизируют код под экосистему CUDA. Компания Т-Платформы планирует дальнейшее расширение линейки вычислительных модулей с использованием ускорителей для повышения производительности и энергоэффективности приложений.

Сверхэффективная система памяти для запуска моделирования сложных проблем и для приложений с интенсивным обменом данными

На модулях V205 расположено на 50% больше слотов DIMM, чем в предыдущем поколении систем T-Blade 1.1a на базе процессора AMD Opteron™. Конфигурации с 8-ядерными процессорами AMD Opteron™ 6220 могут оснащаться 16 ГБ памяти DDR3 в расчете на каждое ядро, а 16-ядерные Opteron™ 6276 – по 8 ГБ на ядро. Кроме того, контроллер памяти DDR3 процессора AMD Opteron™серии 6200 поддерживает частоту 1600 МГц на четырех каналах памяти одновременно, что значительно повышает пропускную способность при операциях чтения и записи из оперативной памяти.

Локальные диски для приложений

Вычислительные модули обоих типов могут быть оснащены двумя жесткими или твердотельными дисками формата 2,5" для хранения промежуточных данных с общим объемом дискового пространства до двух терабайт. Поддерживаются базовые уровни RAID 0/1/10 и возможность бездисковой загрузки с использованием накопителя USB и протоколов iSCSI или PXE.

Уникальный дизайн системной платы

Двухпроцессорная системная плата вычислительного модуля V205 является оригинальным решением компании Т-Платформы. Плата с 16 разъемами DIMM имеет уменьшенную высоту для установки в шасси 5U, что дает заказчикам и партнерам возможность выбора и дифференциации. Кроме того данная системная плата может поставляться в версии под установку в «twin»-серверах.

Основные характеристики узлов V205

Вычислительные узлы v205 обладают следующими основными характеристиками:

Поддерживаемые микропроцессоры и ускорители:

- *AMD Opteron™ серий 6100 и 6200 (4-8-12 и 16-ядерные версии).*
- *NVIDIA® Tesla™ M2075 и M2090 (только для версии V205F).*

Подсистема памяти:

- *До 16 модулей DDR3 RDIMM ECC 1066/1333/1600 МГц (до 256 ГБ на узел);*

Локальная дисковая система:

- *До двух дисков SATA 2.0 (3Гб/с) холодной замены формата 2,5" на узел.*

Слот расширения

- *1 слот PCIe x16 Gen 2.0 для установки адаптера формата Low Profile MD2 (только для версии V205S).*

Интегрированные сети и интерконнект

- *Двухпортовый контроллер GbE (RJ45) и опциональный контроллер QDR Infiniband / 10GbE VPI (один порт QSFP).*

1.2 Позиционирование шасси V5000

Вычислительные модули V205S/F разработаны для шасси V5000 —корпуса в исполнении 5U с оптимизированным энергопотреблением и централизованной подсистемой мониторинга (Изображения 1 и 2).

Шасси V5000 не имеет встроенных коммутаторов сети и интерконнекта, что в комбинации с вычислительными модулями, оборудованными QSFP-портами, для создания вычислительных кластеров с количеством узлов до 648. Использование внешних edge-коммутаторов Infiniband является привлекательным по цене и эффективности решением, позволяющим избежать эффекта «переподписки» портов.

Данное высокофункциональное шасси с избыточными компонентами горячей замены позиционируется как "scale out"- решение для широкого круга HPC-пользователей и допускает смешанную установку вычислительных модулей разных типов.

Изображение 1. Передняя панель шасси V-Class

Изображение 2. Тыльная сторона шасси V-Class

Основные характеристики шасси V5000*

Исполнение

- *5U, для установки в стандартные девятнадцатидюймовые стойки с глубиной не менее 1070 мм.*

Максимальное количество устанавливаемых узлов горячей замены

- *10 двухпроцессорных вычислительных модулей с литерой S.*
- *5 двухпроцессорных вычислительных модулей с литерой F с ускорителями GPU.*
- *Возможность установки смешанных конфигураций вычислительных модулей.*
- *Возможность установки смешанных конфигураций вычислительных модулей*

Системное управление

- *Модуль холодной замены в исполнении 1U с контрольной панелью системы и встроенным коммутатором Fast Ethernet с двумя внешними портами GbE для консолидации мониторинга и управления шасси и узлами.*
- *Поддержка iKVM, Remote Media, Serial over LAN, IPMI2.0 over LAN.*

Система охлаждения

- *3 модуля воздушного охлаждения с избыточностью N+1, с функцией «горячей» замены.*

Система электропитания

- *3 или 4 блока питания 1600 Вт, с избыточностью N+1, с функцией «горячей» замены.*
- *Блоки питания 80Plus Platinum (94%).*

Пиковое энергопотребление:

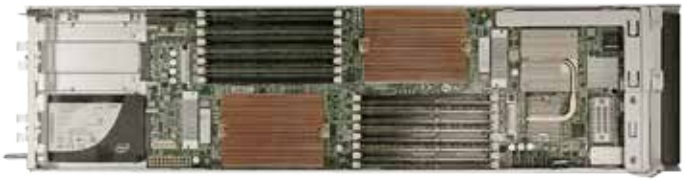
- *4700 Вт (предварительная информация для максимальной конфигурации на базе 5 модулей F с 10 ускорителями GPU NVIDIA® Tesla™ M2090).*

* Дополнительная информация по шасси V5000 изложена в документе под названием «Обзор шасси V5000».

2. Детальный обзор вычислительных модулей V205S и V205F

Шасси V5000 поддерживает два типа вычислительных модулей на базе процессоров AMD Opteron™:

- *V205S является узлом x86 в стандартном корпусе (Изображение 3).*
- *V205F является узлом x86 с установленным ускорителем GPU в корпусе двойной толщины. Дополнительное пространство внутри вычислительного модуля требуется для обеспечения необходимого обдува ускорителя NVIDIA® Tesla™ серии M (Изображение 4).*



Изображение 3.
Вычислительный модуль стандартной толщины V205S с одним SSD (вид сверху, без верхней панели).

Изображение 4.
Вычислительный модуль двойной толщины V205F с ускорителем NVIDIA Tesla M2090 (вид сверху, без верхней панели)

Шасси V5000 поддерживает единообразные или смешанные конфигурации узлов S и F вне зависимости от конфигурации и производителя/типа микропроцессоров. Заказчики могут устанавливать различные типы вычислительных модулей в произвольном порядке. В пустые отсеки в обязательном порядке устанавливаются механические заглушки (blank panels), обеспечивающие корректный воздушный поток внутри шасси. Система управления определяет присутствие и тип модуля в отсеках и обеспечивает актуальное отображение системы в графическом интерфейсе и на контрольной панели шасси. Поддерживается «горячая» замена узлов.

Дополнительная информация о порядке установки вычислительных модулей приведена в документе «Обзор шасси V5000».

2.1 Общая конструкция вычислительных модулей V205S и V205F

В отличие от большинства систем «twin»-класса, оба модуля поставляются в закрытых корпусах со съемной верхней панелью. Вычислительные модули V205 не имеют вентиляторов внутри узла и практически не содержат кабельных соединений, что повышает надежность системы. Исключение составляет лишь кабель дополнительного питания ускорителя NVIDIA Tesla M20xx, обязательный для использования в узле V205F.

Оба корпуса вычислительных модулей оборудованы прямоугольными отверстиями на входе и сотовыми отверстиями на выходе, обеспечивающими оптимальный воздушный поток, нагнетаемый модулями охлаждения шасси. На обоих корпусах установлена панель портов ввода-вывода (Изображение 5). В модуле V205S также предусмотрен кронштейн для установки одного адаптера расширения стандарта PCIe x16 LP MD2. В зависимости от типа корпуса поддерживается установка адаптера PCIe x16, например, дополнительного контроллера Infiniband, или один ускоритель GPU с TDP до 250 Вт.

Детальная информация о портах и индикаторах модулей приведена в секции 2.2 данного документа.

На внутренней стороне модуля установлены специализированный разъем и направляющая для сопряжения вычислительного модуля с объединительной платой шасси.

Изображение 5.
Вычислительные модули V205S (слева) и V205F (справа).

В корпус вычислительных модулей S и F устанавливается унифицированная двухпроцессорная системная плата с пассивной платой-разъемом CardEdge для сопряжения с шасси (Изображение 6). Системная плата оснащена двумя разъемами SATA для прямого подсоединения двух дисков SATA 2.0 «холодной» замены формата 2,5".

Изображение 6. Плата-разъем CardEdge, установленная на V205S.

Вычислительный модуль V205F содержит направляющие потока воздуха для охлаждения ускорителя NVIDIA® Tesla™ серии M (Изображение 7).

Изображение 7.
Узел V205F с установленным ускорителем NVIDIA Tesla M2090.

2.2 Обзор системной платы V205

Системная плата V205 (Изображение 8) относится к платформе AMD G34 и поддерживает новейшие 4-8-12- и 16-ядерные процессоры AMD Opteron™ серии 6200, и 8-12-ядерные процессоры серии 6100 с TDP до 115 Вт для системы V-Class.

Плата содержит специализированную плату-разъем CardEdge для коммутации входного напряжения и сигналов индикации и управления. Системная плата V205 имеет уменьшенную на несколько миллиметров высоту (по сравнению с большинством сторонних плат G34, оборудованных 16 разъемами DIMM, доступными на момент выхода платы на рынок), что позволяет установить ее в отсек шасси высотой 4U.

Изображение 8. Системная плата V205-S1A

Габариты системной платы V205:

- *длина: 506 мм (20");*
- *ширина: 165 мм (6,5").*

Системная плата V205 поставляется в версиях B (блейд) для установки в шасси V5000 и S (сервер) с выделенным портом управления Fast Ethernet и дополнительными разъемами SATA и электропитания.

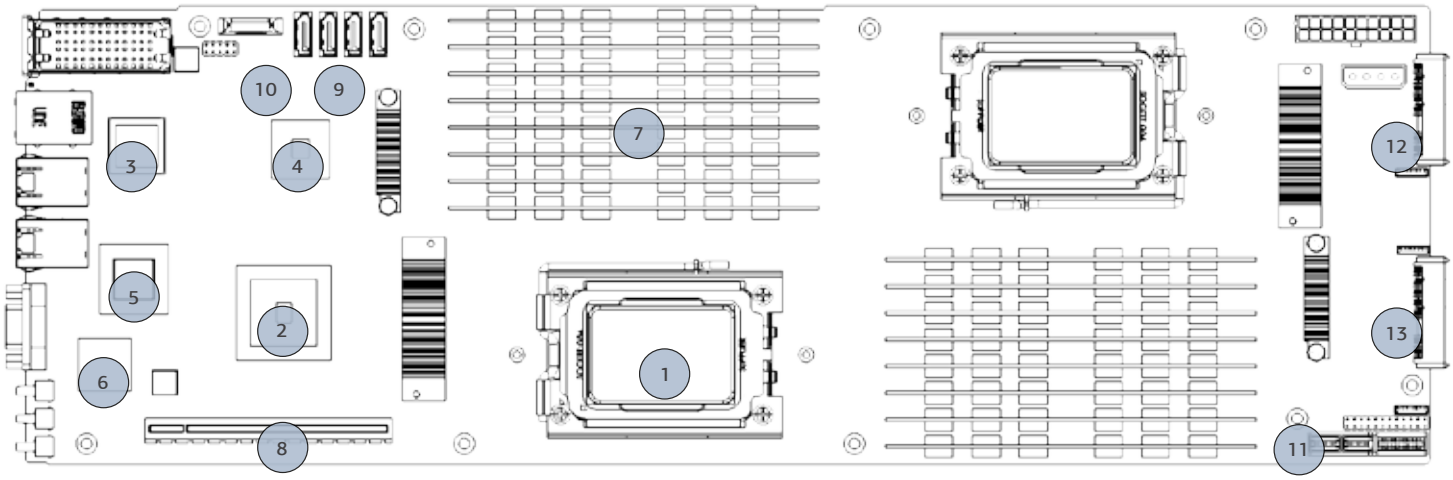
Версия B:

- *V205-B1A – SKU с интегрированным контроллером Infiniband для установки в шасси V5000.*
- *V205-B0A – SKU без интегрированного контроллера Infiniband для установки в шасси V5000.*

Версия S:

- *V205-S1A – SKU с интегрированным контроллером Infiniband для установки в твин-сервер.*
- *V205-S0A – SKU без интегрированного контроллера Infiniband для установки в твин-сервер. **

* За дополнительной информацией по доступности плат в версии S просьба обращаться в отдел продаж компании Т-Платформы.



Изображение 9.
Расположение основных компонентов на плате V205 (схема инженерного образца со всеми возможными опциями, портами и коннекторами)

Разъемы для микропроцессора

1. Гнездо G34 для ЦП AMD Opteron с термопакетом до 115 Вт TDP

Интегрированные контроллеры

2. Северный мост AMD SR5670
3. Южный мост AMD SP5100
4. Однопортовый контроллер IB Mellanox ConnectX2 (опция)
5. Двухпортовый GbE-контроллер Intel 82580DP
6. Микросхема BMC/VGA ASPEED 2050

Разъемы расширения

7. 16 слотов DIMM DDR3 (8 на процессор)
8. Один слот PCIe x16 Gen 2
9. Четыре разъема SATA 2.0 (опция)
10. Один планарный порт USB2.0

Специализированные разъемы

11. Разъем CardEdge для сопряжения с Midplane
12. 12 Разъем SATA 1
13. Разъем SATA 2

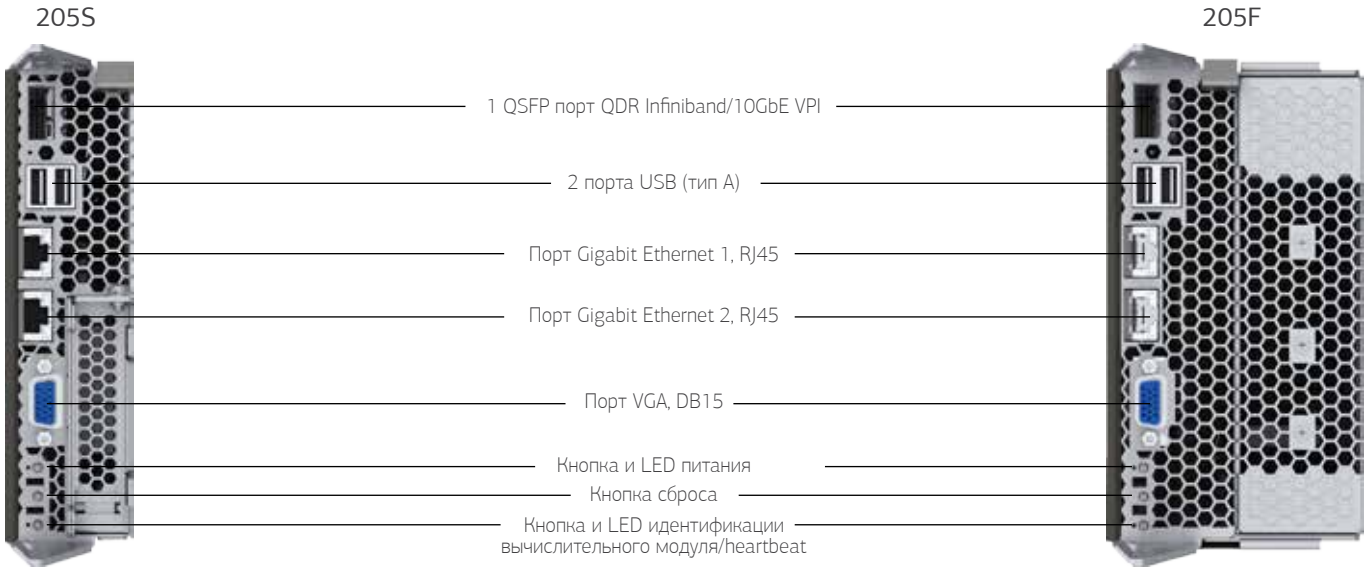


Таблица 1. Панели ввода-вывода вычислительных модулей V205S и V205F

2.3 Обзор платформы AMD 6000

Процессоры AMD Opteron® серии 6000 построены на архитектуре Direct Connect 2.0, которая включает в себя технологию HyperTransport™ и интегрированные на кристаллах процессоров контролеров памяти.

Каждый серверный процессор содержит до четырех соединений «точка-точка» HyperTransport 3.0 для объединения процессоров и устройств ввода-вывода в единую масштабируемую систему.

Интегрированные контроллеры памяти в процессорах AMD Opteron™ серии 6200 поддерживают частоты 1066/1333/1600 МГц для плат с двумя слотами DDR3 DIMM на каждый канал памяти.

Корпорация AMD первой вышла на рынок серверов стандартной архитектуры с платформой, содержащей интегрированные на кристаллах процессоров контроллеры памяти, отказавшись от использования единой процессорной шины. В определенной степени, это повлияло на процесс трансформации серверных систем x86 в доминирующую серверную архитектуру.

2.3.1 Процессор AMD Opteron ™ серии 6100

Процессор AMD Opteron™ серии 6100, представленный весной 2010 года, является первым массовым серверным процессором с двенадцатью ядрами. Основанная на архитектуре «Magny-Cours», серия 6100 устанавливается в 2-4 сокетные платы, устраняя искусственную ценовую разницу, исторически разделявшую два сегмента.

Микропроцессор оборудован четырехканальным контроллером памяти DDR3 и поддерживает скорости DIMM до 1333 МГц. Производимый по технологии 45нм, данный процессор не достигает частотных характеристик процессоров Intel Xeon серии X5600, однако количество доступных ядер, высокая пропускная способность памяти и сравнительно низкий тепловой конверт «Magny-Cours» позволяет с большим успехом использовать его в тех приложениях HPC, которые мало зависят от высокой частоты микропроцессоров.

Стратегия долгосрочных серверных платформ AMD обеспечивает защиту инвестиций, поддерживая новейшее поколение микропроцессоров — AMD Opteron™ серии 6200.

Модель	Частота	Кэш L2	Кэш L3	Кол-во ядер	Частота HT	Мощность, TDP	Тип гнезда
61KS	2,0 ГГц	8 x 512 КБ	2x6M	8	6400 MT/c	115 Вт	Socket G34
61QS	2,3 ГГц						
6128	2,0 ГГц						
6134	2,3 ГГц						
6136	2,4 ГГц						
6140	2,6 ГГц						
6124 HE	1,8 ГГц					85 Вт	
6128 HE	2,0 ГГц						
6132 HE	2,2 ГГц						
6168	1,9 ГГц	12 x 512 КБ		12		115 Вт	
6172	2,1 ГГц						
6174	2,2 ГГц						
6176	2,3 ГГц						
6164 HE	1,7 ГГц					85 Вт	
6166 HE	1,8 ГГц						

Таблица 2. Процессоры AMD Opteron серии 6100, поддерживаемые V205.

2.3.2 Процессор AMD Opteron™ серии 6200

Новейший процессор «Interlagos», представленный рынку в ноябре 2011 года, основан на принципиально новом дизайне двухядерного модуля под названием «Bulldozer». Каждый процессор содержит два кристалла с несколькими двухядерными модулями, соединенных интерфейсами HyperTransport 3.0.

Новый 32нм процессор обладает более высокой производительностью, масштабируемостью и рядом новых характеристик:

- *4-8-12 и 16-ядерные версии ЦП;*
- *увеличенные пропускная способность и частота работы памяти до 1600 МГц на канал;*
- *4 интерфейса HyperTransport 3.0 со скоростью до 6,4 Гигатрансферов в секунду;*
- *256-разрядный (или 2x128 разрядных) блок(а) обработки плавающей точки FlexFP с поддержкой векторных расширений (AVX) для ускорения вычислений в научных, мультимедийных и финансовых приложениях и в целом параллелизма выполнения;*
- *новое операционное состояние TurboCore 1 для увеличения частоты всех ядер на 500 МГц – 1 ГГц;*
- *режим TurboCore 2, который позволяет увеличить частоты половины ядер на 1ГГц;*
- *агрессивное выключение транзисторной логики для снижения энергопотребления;*
- *минимальный режим P-State – 500 МГц;*
- *улучшенные режим C1E и новое состояние C6, в котором данные кэшпамяти ЦПУ сохраняются в системной памяти, а процессор в состоянии простоя входит в режим глубокого сна;*
- *технология TDP Capping позволяет задавать верхнюю границу теплового конверта ЦПУ шагом в 1 Вт, что помогает увеличить вычислительную плотность системы в стойке или же не выходить за рамки допустимого энергопотребления.*

2.3.3 Технология HyperTransport 3.0

Технология HyperTransport 3.0 является высокоскоростным интерфейсом точка-точка (6,4ГТ/с) с низкими коммуникационными задержками. Технология HT функционирует как в межпроцессорном когерентном, так и в периферийном режимах. Технология позволяет эффективно использовать подсистему памяти в мультипроцессорных системах, используя арбитры и фильтры межпроцессорного обмена, такие, как HT Assist™, снижающие количество межпроцессорного трафика с синхронизированными копиями данных кэш-памяти процессоров.

Модель	Базовая частота / частота NB	Режим Turbo CORE P1	Режим Turbo CORE P0	Кэш L2	Кэш L3	Кол-во ядер	Частота HT	Мощность, TDP	Тип гнезда
6204	3,3/2	NA	NA	2x2M	16M	4	6400 MT/c	115 Вт	Socket G34
6276	2,3/2	2,6 ГГц	3,2 ГГц	8x2M	16M	16	6400 MT/c	115 Вт	
6274	2,2/2	2,5 ГГц	3,1 ГГц						
6272	2,1/2	2,4 ГГц	3 ГГц						
6238	2,6/2	2,9 ГГц	3,2 ГГц	6x2M		12			
6234	2,4/2	2,7 ГГц	3 ГГц						
6220	3,0/2	3,3 ГГц	3,6 ГГц	4x2M		8			
6212	2,6/2	2,9 ГГц	3,2 ГГц						
6262 HE	1,6/1,8	2,1 ГГц	2,9 ГГц	8x2M	16M	16	6400 MT/c	85 Вт	

Таблица 3. Процессоры AMD Opteron серии 6200, поддерживаемые V205

2.3.4 Обзор набора микросхем AMD

Системная плата V205 основана на контроллерах AMD SR5670 и SP5100, поддерживающих широкую номенклатуру процессоров AMD различных классов. Производимый с начала 2010 года, чипсет является частью стратегии AMD Stable platform, направленной на создание унифицированных платформ с длинным жизненным циклом.

AMD SR5670 является контроллером PCI Express 2.0, подключаемым к процессору AMD Opteron™ через интерфейс HyperTransport 3.0 на скорости 5.2ГТ/с, поддерживающим 30 линий PCI Express 2.0.

Выбор контроллера SR5670 не случаен. SR5670 является наиболее сбалансированным решением AMD для создания вычислительного узла V-Class, обеспечивающим как адекватные возможности для ввода-вывода, так и низкое тепловыделение (15.4 Вт против 18 Вт на 42-линейном SR5690. SR5670 использует интерфейс x4 A-Link для подсоединения периферийного контроллера SP5100.

AMD SP5100 обеспечивает интерфейс 3Гб/с для подключения до 6 устройств Serial ATA и поддержку базовых функций SW RAID, задействуя стек Promise RAID. Контроллер так же поддерживает до двенадцати портов USB 2.0, шину PCI 2.3 и интерфейсы LPC/SPI.

2.3.5 Новые технологии электросбережения AMD-P в ЦП AMD Opteron™ 6200

Технология «AMD-P» является общим названием, объединяющим большинство технологий управления энергопотреблением, используемых в процессорах AMD Opteron™ серий 6100 и 6200. В «AMD-P» можно выделить три основных нововведения:

- *TDP PowerCapping позволяет установить верхний лимит TDP с шагом в 1 Вт. Первоначально срабатывает схема агрессивного отключения незадействованной транзисторной логики, и уже затем меняется режим p-State. Таким образом, заказчики могут покупать наиболее производительные процессоры и точно задавать их максимальное тепловыделение для установки систем в средах, ограниченных по пространству и электропитанию.*
- *C1E является режимом управления энергопотреблением, при котором в процессоре сокращается энергопотребление не только ядер, но и контроллера памяти и других элементов. Интерфейсы HyperTransport™ так же могут вводиться в состояние простоя.*
- *C6 является новым режимом энергосбережения, когда оба ядра в каждом модуле «Bulldozer» находятся в состоянии простоя, а данные кэш-памяти временно хранятся в модулях DRAM. Технология позволяет экономить до 85% энергии за счет отключения части транзисторной базы микропроцессора.*

2.3.6 Технология FlexFP с поддержкой расширений AVX

С момента выпуска процессора Opteron™ серии 6200, AMD стала первой компанией, выпустившей серверный процессор с поддержкой набора инструкций Intel Advanced Vector eXtensions (AVX), 256-рязрядного блока обработки плавающей точки, а также наборов инструкций SSE4.1, SSE4.2, AES, CLMUL и 128-разрядных наборов инструкций SSE, включая операции FMAC, XOP, FMA4 и CVT16.

В набор инструкций AVX добавлено 12 новых инструкций и расширен размер регистров XMM со 128 до 256 разрядов. С помощью новейших компиляторов пользователи могут удваивать количество исполняемых приложениями операций. Также ускоряется исполнение инструкций SIMD, задействованных во многих научных и мультимедийных приложениях.

Главными преимуществами микроархитектуры сопроцессора « FlexFP» являются ее эффективность и гибкость. Пользователи получают 256-разрядную поддержку AVX для оптимизированного ПО, а так же высокую производительность в «классических» приложениях, не использующих расширенные векторные инструкции AVX. Технология FlexFP поддерживает как режим с двумя 128-разрядными блоками исполнения, выделенными для каждого целочисленного ядра, так и общий 256-разрядный режим работы сопроцессора, обслуживающий сразу два целочисленных ядра. Это позволяет уменьшить транзисторную базу ЦПУ и энергопотребление, требуемые для реализации редко используемых функций.

3. Подсистема памяти DDR3

Четырехканальный контроллер памяти DDR3 интегрирован непосредственно на кристалл микропроцессора AMD Opteron™, что снижает задержки в межпроцессорном обмене данными и позволяет масштабировать пропускную способность путем добавления в систему дополнительных или более производительных ЦПУ.

Системная плата V205 использует дизайн памяти «single stripe» для снижения задержек и поддержки высокочастотных DIMM. Также поддерживается режим RAS «Online Spare».

Регистровые модули DIMM

Регистровые модули DIMM обеспечивают увеличение поддерживаемого объема памяти, поскольку контроллер памяти управляет сигналами адресации и команд только для чипа регистрации, что сокращает электрическую нагрузку на контроллер.

В вычислительных узлах V205 может устанавливаться по два регистровых DIMM на каждый из четырех каналов памяти, обеспечивая общий объем в 256ГБ оперативной памяти на узел. Компанией Т-Платформы оттестированы регистровые модули DIMM с коррекцией ошибок ECC на частотах до 1600 МГц включительно (Таблица 4).

Частота работы оперативной памяти зависит от типа модулей, конфигурации установленной памяти и модели используемых микропроцессоров. Поддерживаются модули RDIMM с ECC объемом 2, 4 и 8ГБ.

Заказчики также могут использовать модули DDR3 RDIMM с пониженным напряжением 1.35В (LV), улучшающие тепловые характеристики узлов и снижающие общее энергопотребление системы.

	Регистровые DIMM
Частоты	1066, 1333 и 1600 МГц
Количество рангов	1, 2 или 4
Объем DIMM	1, 2, 4, 8 или 16 ГБ*
Максимальное кол-во DIMM на канал	2 (для V205)
Напряжение	1,5 В и 1,35 В
Технология DRAM	x4 или x8
Температурный датчик	Да
ECC	Да
Advanced ECC	Только x4 DIMM
Address Parity	Да

Таблица 4. Поддерживаемая V205 память DDR.

	Ранг DIMM 1	Ранг DIMM 2	Макс частота, 1.5V DIMMs	Макс частота, 1.35V DIMM	Макс ГБ/ канал
RDIMM	1R или 2R	Пустой	1600 МГц	1333 МГц	8 ГБ
	1R	1R	1600 МГц	1333 МГц	8 ГБ
	1R или 2R	2R	1333 МГц	1333 МГц	16 ГБ
	4R	Пустой	1333 МГц	1066 МГц	16 ГБ
	4R	1R, 2R или 4R	1066 МГц	800 МГц	32 ГБ
LR-DIMM	4R	Пустой	1600 МГц	1333 МГц	16 ГБ
	4R	4R	1333 МГц	1333 МГц	32 ГБ

Таблица 5.
Правила установки регистровых DDR3 DIMM для конфигураций с ЦП AMD Opteron серии 6200

RDIMM	1R или 2R	Пустой, 1R или 2R	1333 МГц	1333 МГц	16 ГБ
	4R	Пустой	1333 МГц	1066 МГц	16 ГБ
	4R	1R, 2R или 4R	1066 МГц	800 МГц	32 ГБ

Таблица 6.
Правила установки регистровых DDR3 DIMM для конфигураций с ЦП AMD Opteron серии 6100

* На момент написания документа модули объемом 16 ГБ не тестировались.

4. Использование ускорителей на базе GPU (V205F)

По мере роста приложений, использующих ускорители GPU, увеличивается и количество гетерогенных систем на базе технологии NVIDIA® Tesla™, присутствующих в списке TOP500. Корпорация NVIDIA прикладывает целенаправленные усилия по развитию экосистемы аппаратного обеспечения, средств программирования и отладки, а также приложений для поддержки вычислений на базе видеоускорителей, активно взаимодействуя с OEM-компаниями и производителями ПО.

Являясь OEM-партнером корпорации NVIDIA, компания Т-Платформы разрабатывает и поставляет гетерогенные вычислительные системы, а также предлагает услуги по оптимизации приложений под экосистему NVIDIA CUDA.

В вычислительный модуль V205F устанавливается один ускоритель NVIDIA® Tesla™ серии M.

20-серия Tesla™ более чем в 10 раз превышает производительность четырехядерных x86 процессоров в операциях с двойной точностью, и поддерживает алгоритмы коррекции ошибок памяти ECC. Ускорители Tesla™ M20xx имеют повышенную надежность и возможность более «плотной» интеграции со средствами системного мониторинга и управления электропотреблением.

В сравнении с V205S вычислительные модули V205F, оснащенные ускорителями NVIDIA® Tesla™ M2090, повышают пиковую производительность системы более чем в три раза, улучшая в два раза энергоэффективность вычислений и повышая вычислительную плотность шасси в полтора раза.

Ускоритель NVIDIA® Tesla™ серии M является полнопрофильным адаптером двойной толщины стандарта PCI Express® второго поколения, основанным на микроархитектуре NVIDIA® Fermi GPU.

Адаптер содержит сам GPU-контроллер, 6 ГБ высокоскоростной памяти GDDR5 и пассивный радиатор, охлаждаемый модулями охлаждения CFM, установленными во фронтальной части шасси (Изображение 11).

Ускорители могут быть сконфигурированы администратором в режимах с включенным ECC для исправления однобитных ошибок и определения двухбитных ошибок, что сократит доступную память до ~5.25 ГБ. Администратор также может снизить тепловой пакет (TDP) ускорителя средствами NVIDIA®.

Ускоритель Tesla M2075 является продуктом среднего класса с TDP 200 Вт; ускоритель M2090 является наиболее производительным продуктом с максимальным тепловыделением до 225 Вт (Таблица 7).

Ускоритель, устанавливаемый в узлы V205F, подключается через дополнительный 8-контактный разъем к системной плате для обеспечения необходимого электропитания.

Изображение 11. Ускоритель NVIDIA® Tesla™ M20x0.

Модель	Основные характеристики	Кол-во поддерживаемых карт | Пиковая производительность	Тепловой конверт TDP
NDIDIA Tesla M2090	Чип GPU • *Количество ядер: 512* • *Частота: 1,3 ГГц* Плата • *Интерфейс PCI Express Gen2 x16* • *Габариты: 111,15 мм x 247 мм, «двойной» слот* Внешние порты • *Отсутствуют* Внутренние порты и разъемы • *Один 6-конт. разъем питания PCI Express* • *Один 8-конт. разъем питания PCI Express* Память • *Частота памяти: 1,85 ГГц* • *Интерфейс: 384-разрядный* • *6 ГБ, GDDR5 SDRAM (5.25 ГБ с ECC)* BIOS • *2Mbit Serial ROM*	1 карта | 665GF DP	<=225 Вт
NVIDIA Tesla M2075	Чип GPU • *Количество ядер: 448* • *Частота: 1,15 ГГц* Плата • *Интерфейс PCI Express Gen2 x16* • *Габариты: 111,15 мм x 247 мм, «двойной» слот* Внешние порты • *Отсутствуют* Внутренние порты и разъемы • *Один 6-конт. разъем питания PCI Express* • *Один 8-конт. разъем питания PCI Express* Память • *Частота памяти: 1,566 ГГц* • *6 ГБ, GDDR5 SDRAM (5.25 ГБ с ECC)* BIOS • *2Мбит Serial ROM 2Мбит Serial ROM*	1 карта | 515GF DP	<=200 Вт

Таблица 7. Характеристики ускорителей NVIDIA® Tesla™ серии M.

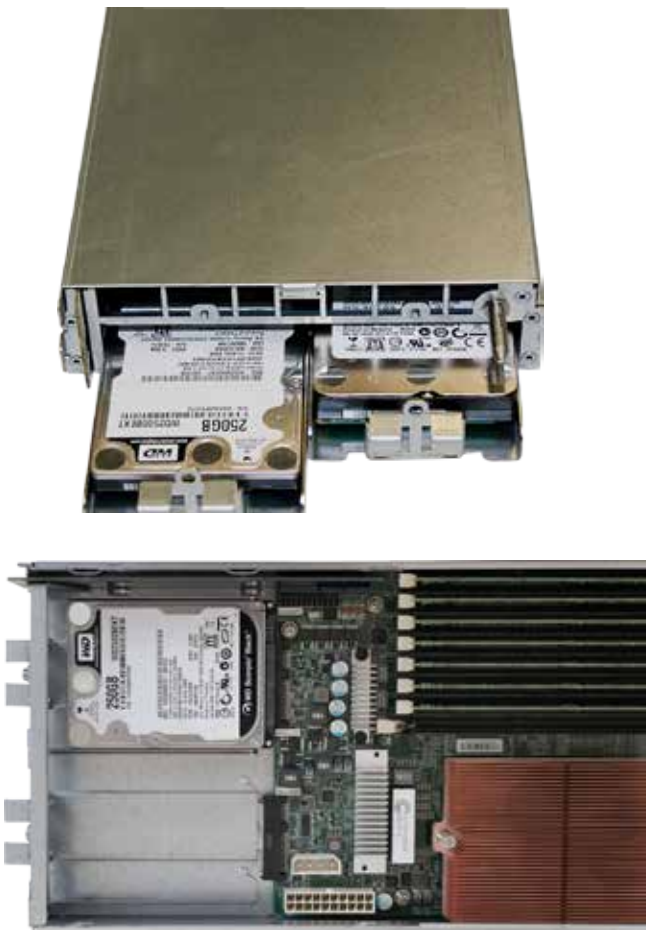
5. Дисковая подсистема

Вычислительные модули V205S и V205F оснащены встроенным контроллером SATA 2.0 и поддерживают до двух дисков с интерфейсом SATA. К заказу так же доступны бездисковые конфигурации узлов, загружающиеся через протоколы iSCSI, PXE и внутренний/внешний USB-носитель.

Микросхема AMD SP5100 поддерживает технологию Promise Software ROMB, реализуя уровни RAID 0/1/10 для двухдисковой подсистемы для ряда ОС семейств Microsoft® и Linux с настройкой параметров через Promise RAID Option ROM и интерфейс WebPAM.

Поддерживаются HDD и SSD формата 2,5 дюйма с толщиной 5, 7, 9,5, 12,5 и 15 мм. Диски монтируются на салазках и подсоединяются напрямую к системной плате, используя разъем SATA CardEdge. Поддерживается функция «холодной» замены дисков, предусматривающая изъятие вычислительного модуля из шасси и последующее извлечение диска без необходимости доступа к системной плате (Изображение 12).

Заказчикам предлагается широкий выбор дисков Seagate, Western Digital и Intel. Доступны как корпоративные диски SATA 2.0 и SATA 3.0* со скоростью вращения 7200RPM и 10000RPM с разным объемом дискового пространства, так и сверхэкономичные твердотельные диски с высочайшим количеством выполняемых IOPS.



Изображение 12.
Узел V205S с дисками холодной замены формата 2.5 дюйма.

Диски семейства Seagate Constellation®, созданные для некритичных приложений «nearline» представлены в двух объемах и обладают показателем в 1,2 миллиона часов наработки на отказ и самым низким в индустрии усредненным потреблением для HDD в 3,11 Вт.

Второе поколение Seagate Constellation®.2 является единственным доступным на рынке 2,5 дюймовых корпоративных дисков семейством с объемом до 1ТБ на момент выхода системы, позволяя хранить почти в два раза больше информации в формате 2,5". Диск объемом в 1ТБ имеет показатель MTBF в 1.4 миллиона часов и среднее энергопотребление в 5.43 Вт. Более высокая плотность записи нового семейства позволила чуть улучшить среднее время поиска данных по сравнению с первым поколением семейства дисков Constellation®.

В отличие от дисков Seagate Constellation®, оптимизированных под некритичные приложения «nearline», и обладающих низкими тепловыми характеристиками, объемом до 1ТБ и удовлетворительной производительностью, диски Western Digital® семейства VelociRaptor® созданы в первую очередь для достижения высокой производительности. Диски WD в формате 2,5" с интерфейсом SATA представлены четырьмя моделями с разным объемом. На 10000 оборотах в минуту они достигают скорость чтения в 3,6 мс (random read) и времени записи 4,2 мс, что приблизительно в два раза быстрее производительности семейств Seagate Constellation®. Высокая производительность, однако, достигается за счет большего энергопотребления и тепловыделения.

Для заказчиков, не стесненных бюджетом и заинтересованных в создании энергоэффективных узлов с высоким показателем IOPS, отличным выбором могут стать твердотельные диски Intel® семейств 320 и 510.

Если диски SATA HDD в среднем достигают показателей 75-100 IOPS на скорости 7200 RPM и 125-150 IOPS на скорости 10000 RPM, устройства Intel® на базе MLC NAND достигают 20000IOPS при чтении и до 8000 IOPS при операциях записи (блоки 4КБ). Задержки чтения/записи сокращаются с уровня ~ 3,6/8,4 мс для HDD, до 65/90 нс, характерных для SSD производства Intel®. Низкое энергопотребление является другим неоспоримым преимуществом технологии твердотельных дисков.

Изготовленные по технологии 25nm диски 2,5" семейства Intel® SSD 320 на базе памяти NAND Flash Multi Level Cell доступны в 6 объемах, обеспечивая постоянную скорость чтения до 270 МБ/с и постоянную скорость записи до 220МБ/с.

Семейство SSD 510 (34nm Intel® NAND Flash Multi Level Cell Memory) состоит из двух моделей в 120ГБ и 250ГБ. Данное семейство имеет гораздо более высокую пропускную способность, чем семейство SSD 320.

Компания Т-Платформы также планирует использовать диски Intel® новой серии SSD 710 (High Endurace Technology eMLC).

Производитель	Семейство	Модель диска	Объем	Скорость вращения, RPM	Кэш, МБ	Интерфейс, ГБ/с*	Тип	Физ. высота, мм
Seagate	Constellation.2	ST925061xNS	250ГБ	7200	64	SATA 6/3/1.5	HDD 2,5"	15
Seagate	Constellation.2	ST950062xNS	500ГБ	7200	64	SATA 6/3/1.5	HDD 2,5"	15
Seagate	Constellation.2	ST9100064xNS	1ТБ	7200	64	SATA 6/3/1.5	HDD 2,5"	15
Seagate	Constellation	ST9160511NS	160ГБ	7200	32	SATA 3/1.5	HDD 2,5"	15
Seagate	Constellation	ST9500530NS	500ГБ	7200	32	SATA 3/1.5	HDD 2,5"	15
WD	VelociRaptor	WD1500BLHX	150ГБ	10000	32	SATA 3/1.5	HDD 2,5"	15
WD	VelociRaptor	WD3000BLHX	300ГБ	10000	32	SATA 3/1.5	HDD 2,5"	15
WD	VelociRaptor	WD4500BLHX	450ГБ	10000	32	SATA 6/3/1.5	HDD 2,5"	15
WD	VelociRaptor	WD6000BLHX	600ГБ	10000	32	SATA 6/3/1.5	HDD 2,5"	15
Intel	SSD 320	Multiple models	80ГБ	NA	NA	SATA 3/1.5	SSD 2,5"	7/9.5
Intel	SSD 320	Multiple models	120ГБ	NA	NA	SATA 3/1.5	SSD 2,5"	7/9.5
Intel	SSD 320	Multiple models	160ГБ	NA	NA	SATA 3/1.5	SSD 2,5"	7/9.5
Intel	SSD 320	Multiple models	300ГБ	NA	NA	SATA 3/1.5	SSD 2,5"	7/9.5
Intel	SSD 320	Multiple models	600ГБ	NA	NA	SATA 3/1.5	SSD 2,5"	7/9.5
Intel	SSD510	SSDSC-2MH120A2XX	120ГБ	NA	NA	SATA 6/3/1.5	SSD 2,5"	9.5
Intel	SSD510	SSDSC-2MH250A2XX	250ГБ	NA	NA	SATA 6/3/1.5	SSD 2,5"	9.5

Таблица 8. Список дисков, доступных для заказа с узлами V205 по состоянию на ноябрь 2011 года.

* Все HDD и SSD устройства SATA 3.0, установленные в V205, функционируют в режиме 3Гб/с (ограничение интегрированного в AMD SP5100 контроллера SATA)

6. Сетевая инфраструктура

6.1 Интерконнекты QDR Infiniband и 10GbE Ethernet VPI

Системная плата V205 поставляется с опциональным контроллером Mellanox® ConnectX-2 VPI®. Интерфейс виртуального протокола (Virtual Protocol Interface) обеспечивает поддержку как технологии QDR Infiniband, так и 10Gb Ethernet через единый порт QSFP; включенный по умолчанию режим IB может быть изменен на «автоматический» или на «только Ethernet».

Порт QSFP может быть подключен к пассивным медным кабелям Infiniband, активным медным кабелям, и к оптоволокну; для унификации кабельной инфраструктуры можно задействовать гибридный кабель «QSFP в SFP+», производимый компанией Mellanox®.

Характеристики контроллера Mellanox® ConnectX-2:

Контроллер обладает следующими основными характеристиками:

- *Протокол виртуального интерконнекта (VPI).*
- *Одночиповая архитектура.*
- *Интегрированный контроллер SerDes.*
- *Отсутствие необходимости использования локальной памяти.*
- *Задержки MPI в 1мкс.*
- *Выбор 10, 20, или 40Гб/с Infiniband или 10GbE.*
- *Интерфейс PCI Express 2.0 (до 5Гт/с).*
- *Функция разгрузки ЦП от операций передачи данных.*
- *Поддержка моментального чтения-записи (Atomic operations).*
- *Поддержка до 16 миллионов каналов ввода-вывода.*
- *Аппаратная поддержка QoS congestion control.*
- *Аппаратная поддержка виртуализации ввода-вывода.*
- *Аппаратная поддержка операций TCP/UDP/IP (stateless offload).*
- *Инкапсуляция протокола Fibre Channel (FCoIB или FCoE).*

Значимость технологии Infiniband

За последние годы технология Infiniband стала активно использоваться в сегменте HPC, в корпоративных ЦОД и в облачных вычислениях. Обеспечивая низкие коммуникационные задержки, высокую пропускную способность, низкую нагрузку на центральные процессоры и режим Remote Direct Memory Access (RDMA), технология Infiniband стала массовым интерконнектом, пришедшим на замену проприетарных или низкопроизводительных решений. Архитектура Infiniband является коммуникационной фабрикой открытого стандарта, обеспечивающей подключение и масштабируемость десятков тысяч вычислительных узлов или узлов хранения данных, и эффективное использование ресурсов.

Широкая совместимость, обеспечиваемая альянсом компаний Open Fabric, превращает Infiniband в экономичное, энергоэффективное и унифицированное решение для разных топологий сетей с большим количеством клиентов и протоколов.

Усовершенствованные механизмы консолидации сетевых фабрик и типов узлов для вычисления, управления и хранения данных позволяют использовать Infiniband в виде унифицированного интерконнекта без потери производительности, сокращая, таким образом, капитальные и операционные затраты.

Технология CORE-Direct™ обеспечивает разгрузку ЦП от таких коллективных операций MPI, как широкое вещание (broadcasting), сборка (gathering) и глобальная синхронизация коммуникационных примитивов (communication routines).

MPI-кластеры на базе Infiniband

Протокол MPI обеспечивает коммуникационный сервис для распределённых процессов, запускаемых в приложениях HPC. MPI является стандартным и доминирующим слоем в коммуникационных средах современных параллельных кластерных систем, производительность которых сильно зависит от величины задержек, возникающих при межузловом обмене. Исторически Infiniband имеет сверхнизкие задержки между приложениями и высокую пропускную способность вкупе с низкой нагрузкой на ЦП.

Эффективность архитектуры Infiniband основана на технологии передачи сообщений по каналам, уменьшенном количестве копий в памяти всех передаваемых данных и, в отличие от TCP/IP, на механизме избегания обработки трафика стеком операционной системы. Существует несколько реализаций MPI на базе технологии Infiniband, позволяющих полностью реализовать преимущества последней.

IB Subnet Manager вместе с Subnet Administrator обеспечивают сравнительно простую настройку интерконнекта и разных топологий, предоставляя администраторам средства развертывания, мониторинга и диагностики фабрики Infiniband.

Ускорение для систем хранения данных

Сегодня для обеспечения высокой производительности СХД с блочным и файловым методами доступа можно использовать протокол Infiniband RDMA. Поддержка инкапсуляции пакетов Fibre Channel по Infiniband (FCoIB) позволяет организовать доступ к высокопроизводительным СХД корпоративного уровня. По мере того, как технология становится зрелой, крупные вычислительные центры по всему миру начинают создавать мультипетабайтные хранилища первого уровня, используя инфраструктуру QDR или FDR Infiniband.

Протоколы верхнего уровня

Набор протоколов верхнего уровня (ULP), разрабатываемых в составе OFED, позволяет многим популярным приложениям использовать Infiniband. Богатый набор протоколов ULP обеспечиваемый альянсом Open Fabrics, включает в себя следующие протоколы:

- ***MPI*** *– MPI ULP для высокопроизводительных кластеров с полной поддержкой функциональных команд MPI.*
- ***IPoIB*** *– IP over Infiniband. Позволяет приложениям в сети Infiniband, обмениваться сообщениями TCP/IP по сети Infiniband.*
- ***iSER*** *– Расширения iSCSI для служб RDMA. Протокол iSCSI позволяет подключаться к СХД с блочной архитектурой по стандартным TCP/IP-сетям для консолидации сетевой инфраструктуры на базе IP или для создания хранилищ второго уровня.*
- ***NFS-RDMA*** *– Network File System over RDMA. NFS является популярной сетевой файловой системой, обеспечивающий групповой доступ по стандартным TCP/IP-сетям.*
- *Поддержка* ***Lustre****, параллельной файловой системы, часто применяемой в вычислительных системах компании Т-Платформы, позволяет организовать параллельный доступ вычислительных узлов к данным. Возможность использования архитектуры канального ввода-вывода (Infiniband's Channel I/O architecture), позволяет каждому узлу установить независимый, защищенный канал с серверами Lustre Metadata Servers (MDS) и ассоциированными серверами и целями хранения объектов Object Storage Servers и Targets (OSS, OST).*

Набор протоколов верхнего уровня обеспечивает приложения интерфейсами для подключения к сетевым, вычислительным службам, службам хранения и др. При этом любая из этих служб использует лишь одну базовую сеть — Infiniband.

6.2 Сеть Gigabit Ethernet

Узлы V205 поставляются с интегрированным двухпортовым контроллером Gigabit Ethernet Intel 82580DB (LOM) — небольшой энергоэффективной микросхемой, обеспечивающей подсоединение узла к второстепенным и вспомогательным сетям Ethernet.

Контроллер 82580DB использует интерфейс PCI Express x4 (PCIe v2.0; 2,5Гб/с). Для поддержки трансляции и приема менеджмент-пакетов контроллер также подсоединяется к микросхеме BMC AST2050 на системной плате V205. Порты GbE могут конфигурироваться в BIOS для одновременной передачи менеджмент-пакетов вместе с трафиком Ethernet.

Два порта RJ45 обеспечивают подсоединения 1000BASE-T, с режимами 1Гб/с full duplex, 10/100 Мб/с (full/half duplex).

Полный список функций и характеристик доступен в документах Intel® 82580EB/82580DB GbE Controller Feature Software Support Summary и Intel® 82580EB/82580DB Gigabit Ethernet Controller Datasheet на сайте корпорации Intel®.

6.3 Опциональная сеть управления Fast Ethernet

Версии системных плат V205 для установки в twin-серверы могут поставляться с дополнительным выделенным портом Fast Ethernet, расположенным над двумя внешними портами USB.

Версии системных плат V205 для систем V-Class поставляются без выделенного порта управления в силу используемого в шасси V5000 коммутатора управления, консолидирующего мониторинг и управление узлами через внутренние порты вычислительных модулей. Это позволяет значительно сократить количество кабелей 5 категории, используемых для управления вычислительными модулями.

7. Мониторинг и управление уровня вычислительного узла

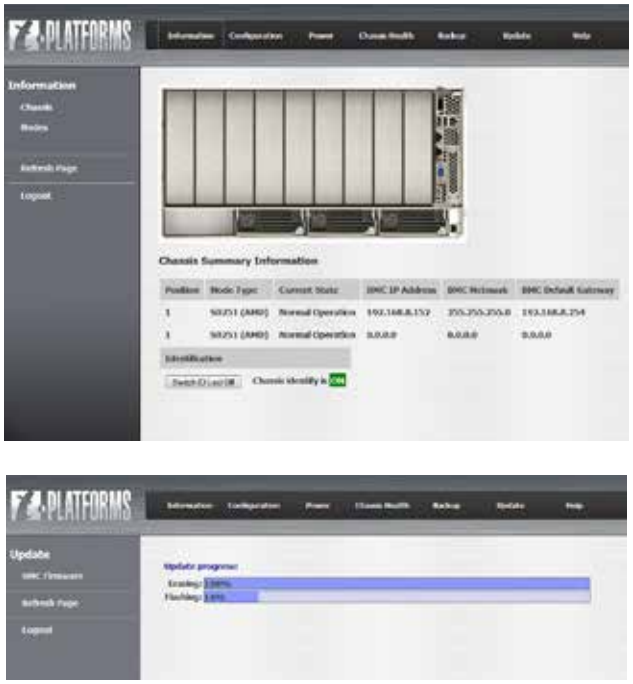
Каждый узел V205 содержит интегрированный контроллер BMC ASPEED 2050, обеспечивающий поддержку интерфейса IPMI 2.0, локальные и удаленные управление и мониторинг через интерфейсы командной строки и веб-сервер.

Защищенный доступ к индивидуальным вычислительным модулям может осуществляться путем прямого подсоединения к внешнему GbE-порту каждого сервера либо централизованно, через единый порт GbE контроллера системного управления (SMC), установленного в шасси V5000.

Шины I2C/IPMB в комплексе с BMC позволяют контролировать интегрированные на системной плате датчики и собирать данные мониторинга. Один из портов GbE на узле может быть активирован исключительно для передачи данных сервисной сети или для смешанного режима передачи сервисной информации и стандартного трафика.

Мониторинг вычислительных модулей позволяет контролировать следующие параметры:

- *температура ЦПУ;*
- *температура северного и южного мостов;*
- *температура модулей памяти DIMM;*
- *температура контроллера Infiniband;*
- *температура регуляторов напряжения ЦПУ и памяти;*
- *напряжение ядер ЦПУ;*
- *напряжение модулей памяти;*
- *основные напряжения на плате (12V, 5VSB, 3.3V, аккумулятор);*
- *Watchdog/NMI.*



Изображение 13. Интерфейс встроенного программного обеспечения IMU.

Управление вычислительным модулем позволяет настраивать следующие параметры:

- *выделение статических IP-адресов каждому BMC (через контроллер управления системой в шасси);*
- *удаленное включение, выключение, цикл и перезагрузка платы;*
- *перезагрузка BMC (холодная и горячая);*
- *удаленное обновление BIOS;*
- *удаленное обновление прошивок BMC;*
- *поддержка iKVM;*
- *поддержка Remote Media;*
- *поддержка Serial over LAN (SOL).*
- *поддержка Remote Media;*
- *поддержка Serial over LAN (SOL).*

Встроенный контроллер системного управления (SMC), установленный в шасси V5000, позволяет обеспечить централизованный сквозной доступ к BMC узлов с помощью интерфейса встроенного ПО IMU (Изображение 13).

Подробная информация о SMC и IMU доступна в документе «Обзор шасси V5000».

8. Поддержка операционных систем

Узлы V205S и V205F поддерживают следующие операционные системы: *

- *SuSE® Linux Enterprise Server 11, Service Pack 1, x86_64;*
- *Red Hat® Enterprise Linux 6 or later (6.1, 6.2) x86_64;*
- *CentOS v6.0, x86_64;*
- *Scientific Linux 6.1, x86_64;*
- *Microsoft® Windows Server® 2008 R2, Service Pack 1 or later, 64-bit.*

Невалидированные ОС с частичной поддержкой:

Ubuntu latest LTS (12.04), x86_64;
Debian v6.0.x, x86_64;
VMWare® ESX 4.0(i);
Citrix® XenServer 6.

9. Заключение

Семейство V-Class

V-Class является новейшей вычислительной платформой для образовательного, научного и коммерческого рынков. Разработанные компанией Т-Платформы кластерные системы на базе V-Class задействуют общепринятые технологии, позволяющие создать гибкий аппаратно-программный комплекс, отвечающий современным требованиям заказчиков. Отсутствие в шасси коммутаторов сетей и интерконнекта позволяет использовать предпочитаемые внешние коммутаторы, а встроенная система управления обеспечивает централизованное управление шасси и узлами.

Системное шасси V5000

Характеристики шасси позволяют заявлять о создании продукта в ценовом диапазоне twin-серверов (с двумя серверными платами, располагаемыми в пространстве 1U) с централизацией управления вычислительными модулями, присущей дорогостоящим блейд-системам корпоративного уровня с более широкой функциональностью.

Кроме того, в отличие от большинства серверов twin-класса, V5000 имеет полную избыточность и «горячую» замену модулей охлаждения, блоков питания.

Шасси T-Blade V поддерживает до 10 двухпроцессорных вычислительных модулей V205S с внешними портами Infiniband и GbE, и допускает установку до 5 вычислительных модулей V205F с ускорителями на базе GPU. В первой половине 2012 года планируется запуск вычислительных модулей на базе процессоров Intel® Xeon® E5 2600 «SandyBridge».

Энергоэффективный дизайн и интегрированная система управления сокращают количество кабельной инфраструктуры, что также делает данное шасси великолепной платформой для создания высокомасштабируемых вычислительных кластеров.

Вычислительные модули V205

Вычислительные модули V205S привлекательны для использования в многопоточных средах, использующих возможности 32 ядер каждого узла. Вычислительный модуль может также поставляться в конфигурации с 8 ядрами на сверхвысокой фиксированной частоте 3,3 ГГц, с увеличенной пропускной способностью системной памяти.

Вычислительный модуль V205F на базе унифицированной с V205S системной платы поддерживает тот же набор центральных процессоров, памяти и дисков, и допускает установку одного ускорителя NVIDIA® Tesla™ серии M.

По состоянию на вторую половину 2011 года платформа AMD 6000 использовалась в половине из 25 наиболее производительных систем списка TOP500, создавая необходимую конкуренцию на рынке серверов стандартной архитектуры. Системы на базе узлов с микроархитектурой «Bulldozer» увеличивают вычислительную плотность, предоставляют в распоряжение пользователей высокоэффективную систему памяти, возможность использования турбочастот до 3,6 ГГц, функционал новейшего 256-разрядного сопроцессора FlexFP с поддержкой AVX и FMA4, и такие технологии гибкого управления питанием, как TDP Capping.

Система V-Class на базе вычислительных модулей V205 доступна для приобретения в России через компанию Т-Платформы или через партнеров компании в ряде других стран.

Актуальная информация по вышедшим и готовящимся к выходу системам доступна на сайте www.t-platforms.ru/products

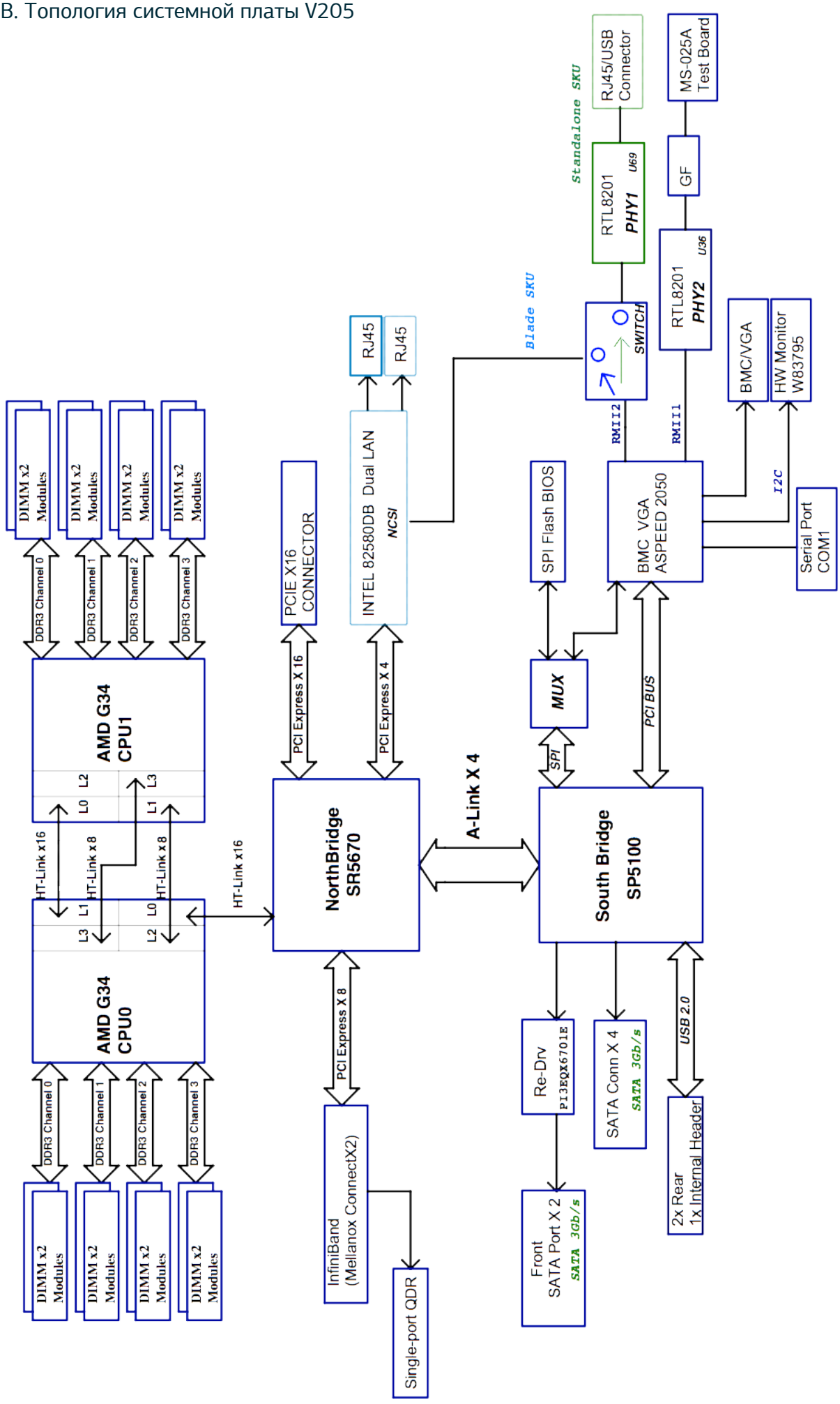
*При использовании ЦП AMD Opteron™ серии 6200 'Interlagos' может потребоваться обновление версии ОС или установка патча ядра для полной реализации и поддержки новой микроархитектуры. Обратитесь в техническую службу компании для получения актуального списка ОС.

10. Приложение.

A. Таблица характеристик систем на базе вычислительных модулей V205S и V205F

Характеристики	Система на базе узла V205S		Система на базе узла V205F	
	Один узел	Одно шасси	Один узел	Одно шасси
Исполнение шасси	5U, для фиксированной установки в стойку			
Исполнение узла	Стандартное (тонкое)		Двойной толщины	
Максимальное количество устанавливаемых узлов	N/A	10	N/A	5
Тип микропроцессора, макс.	AMD Opteron™ 6276, 2,3 ГГц, 16 ядер/16 потоков, TDP 115 Вт			
Максимальное количество гнезд/ядер x86	2/32	20/320	2/32	10/160
Поддержка ускорителей	Не поддерживается		NVIDIA® Tesla™ M2075 или M2090	
Максимальное количество ускорителей	Не поддерживается		1	5
Пиковая производительность DP, GFLOPS,	294.4	2944	959	4795
Пиковая производительность на ватт потребляемой энергии	>0,73GFLOPS/Вт, предварительные данные		1,45GFLOPS/Вт, предварительные данные	
Поддерживаемый тип памяти	Регистровый ECC DDR3 DIMM, 4 канала, 1066, 1333, 1600 МГц			
Количество слотов памяти	16	160	16	80
Максимальный объем оперативной памяти	256 ГБ	2,56 ТБ	256 ГБ	1,28 ТБ
Доступный максимальный объем памяти на ядро	До 8 ГБ (для 16-ядерных моделей ЦП)			
Тип дисковой системы	Локальная, с холодной заменой, SATA 2.0 (3 Гб/с), 2,5" HDD или SSD			
Максимальное количество дисков	2	20	2	10
Доступное дисковое пространство	До 2 ТБ	До 20 ТБ	До 2 ТБ	До 10 ТБ
Набор микросхем вычислительного модуля	Контроллер AMD SR5670; контроллер AMD SP5100			
Тип интерфейса PCI Epress	PCIe x16, Gen.2			
Количество слотов расширения PCI E x16	1 LP MD2	10 LP MD2	1 (только для GPU)	5 (только для GPU)
Порты Ethernet на вычислительных модулях	Два внешних порта GbE на узел (Intel® 82580DB); опциональный интерфейс 10GbE VPI ® через QSFP порт (Mellanox® ConnectX-2)			
Порты QDR Infiniband на вычислительных модулях (опция)	1	10	1	5
Встроенный коммутатор Ethernet	Только встроенная сеть управления Fast Ethernet			
Встроенный коммутатор Infiniband	Не поддерживается			
Встроенная система управления шасси и узлами	Контроллер SMC с коммутатором управления Fast Ethernet и двумя внешними портами GbE; поддержка iKVM и Remote Media; встроенный интерфейс управления IMU			
Тип охлаждения шасси	Воздушное, спереди-назад, 3 модуля со сдвоенными вентиляторами, с функцией горячей замены и избыточностью N+1			
Тип блоков питания в шасси	Три или четыре блока питания 1,6 кВт (@220В) , "80Plus Platinum", с функцией горячей замены и избыточностью N+1			
Пиковое энергопотребление (peak), Вт	N/A	~4200	N/A	~3300
Установившееся энергопотребление (sustained), Вт	N/A	~3800	N/A	~3000
Энергопотребление в состояние простоя (idle), Вт	N/A	~1300	N/A	~1050
Характеристики электросети	1) 208-230 VAC, 50-60Hz, 4 x 8A, с 1 или 3 фазами 2) Часть конфигураций поддерживает подключение к сети с напряжением 110-120 VAC , 50-60Hz, 4 x 16A, с 1 или 3 фазами			
Вес системы без кабелей и рельс, кг	~5,7	~95,35	~8,10	~78,85
Габариты системы с установленными вычислительными модулями, мм	Высота 222,5(5U) x ширина 443 x глубина 868			
Тип поддерживаемых стоек	Стандартные стойки 19 дюймов с глубиной не менее 1070 мм (EIA-310 или более поздний стандарт)			
Сертификация	TBA			

B. Топология системной платы V205



C. Используемые аббревиатуры

- *AVX – Advanced Vector Extensions (расширенные векторные инструкции)*
- *BMC – Baseboard Management Controller (контроллер управления на системной плате)*
- *CUDA – Compute Unified Device Architecture (экосистема NVIDIA® для ускорителей GPU)*
- *DDR3 – Double Data Rate 3 (тип памяти)*
- *DIMM – Dual Inline Memory Module (стандарт модулей памяти)*
- *ECC – Error Checking and Correction (алгоритм детекции и коррекции ошибок памяти)*
- *EIA – Electronic Industries Alliance (бывший альянс компаний по электронным стандартам)*
- *eMLC – enterprise Multi-Level Cell flash technology (технология твердотельных дисков)*
- *FMA – Fused Multiply Add (операция одновременного умножения и сложения в сопроцессоре)*
- *GbE – Gigabit Ethernet*
- *GDDR – Graphics Double Data Rate (тип памяти, применяемый в графических ускорителях)*
- *GPU – Graphics Processing Unit (графический ускоритель)*
- *HDD – Hard Disk Drive (жесткий диск)*
- *HS – Hot Swap (операция «горячей» замены)*
- *IEC – International Electrotechnical Commission (Международная электротехническая комиссия)*
- *IPMB – Intelligent Platform Management Bus (интеллектуальная шина управления платформой)*
- *IPMI – Intelligent Platform Management Interface (интеллектуальный интерфейс управления платформой)*
- *IPoIB – Internet Protocol over Infiniband (инкапсуляция IP-протокола по сети Infiniband)*
- *ISV – Independent Software Vendor (независимый разработчик ПО)*
- *LOM – LAN on Motherboard (интегрированный контроллер сети на системной плате)*
- *MPI – Message Passing Interface (интерфейс передачи сообщений во многих кластерных системах)*
- *NFS – Network File System (сетевая файловая система)*
- *OFED – Open Fabrics Enterprise Distribution (промышленный дистрибутив открытых сетей)*
- *QDR IB – Quad Data Rate Infiniband (протокол InfinBand с передачей данных 40Гб/с)*
- *RAID – Redundant Array of Inexpensive Disks (избыточный набор недорогих дисков)*
- *RDMA – Remote Direct Memory Access (метод удаленного прямого доступа в память)*
- *RU (U) – Rackmount Unit (единица размерности устанавливаемого в стойку оборудования)*
- *SATA (Serial ATA) – Serial Advanced Technology Attachment (последовательный интерфейс подключения дисков)*
- *SIMD – Single Instruction Multiple Data (программные мультимедийные расширения в ЦПУ)*
- *SSD – Solid State Disk (твердотельный диск)*
- *TDP – Thermal Design Power (тепловой диапазон ЦПУ)*
- *ULP – Upper Level Protocol (протокол верхнего уровня)*
- *VPI – Virtual Protocol Interconnect (виртуальный протокол интерконнектов, унифицирующий подключение к адаптерам коммутаторов Infiniband и 10/40 Gigabit Ethernet*

«Т-Платформы»

Москва, Россия, Ленинский проспект, д. 113 / 1, офис В-705
Тел.: +7 (495) 956 54 90
Факс: +7 (495) 956 54 15

tPlatforms GmbH

Woehlerstrasse 42, D-30163, Hannover, Germany
Tel.: +49 (511) 203 885 40
Fax.: +49 (511) 203 885 41

www.t-platforms.ru