

ИНФОРМАЦИОННЫЕ МОДЕЛИ ПРЕДПРИЯТИЯ В РЕАЛИЗАЦИИ ТЕХНОЛОГИИ IN-MEMORY DATA MANAGEMENT.

Гимранов Ринат Дамирович, ОАО «Сургутнефтегаз».

Конференция МГУ «Современные информационные технологии и ИТ-образование».

Москва 14.11.2014

Jim Gray, Microsoft, December 2006:

RAM Locality is King

Tape is Dead

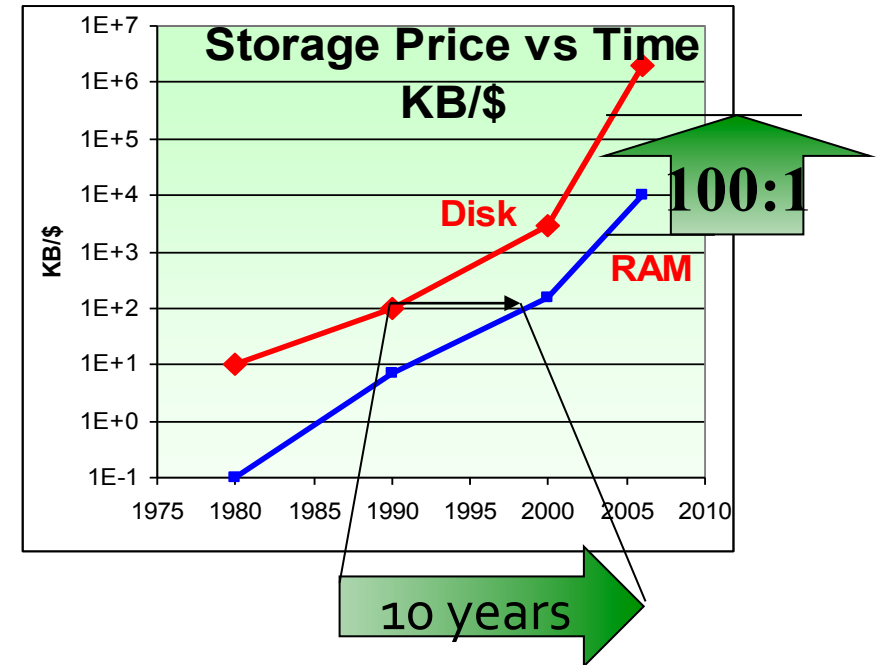
Disk is Tape

Flash is Disk

RAM Locality is King

The cpu mostly waits for RAM

- Flash / Disk are 100,000 ...1,000,000 clocks away from cpu
- RAM is ~100 clocks away unless you have locality (cache).
- If you want 1CPI (clock per instruction) you have to have the data in cache (program cache is “easy”)
- This requires cache conscious data-structures and algorithms sequential (or predictable) access patterns
- Main Memory DB is going to be common.



“The 1970 - 1985 period was a time of intense debate, a myriad of ideas, and considerable upheaval. We predict the next fifteen years will have the same feel.”

“One Size Fits All”: An Idea Whose Time Has Come and Gone

Michael Stonebraker
Computer Science and Artificial
Intelligence Laboratory, M.I.T., and
StreamBase Systems, Inc.
stonebraker@csail.mit.edu

Uğur Çetintemel
Department of Computer Science
Brown University, and
StreamBase Systems, Inc.
ugur@cs.brown.edu

Abstract

*The last 25 years of comm
can be summed up in a single
This phrase refers to the fact
architecture (originally des
business data processing) has
data-centric applications*

of multiple code lines causes various practical problems,
including:

The End of an Architectural Era (It's Time for a Complete Rewrite)

Michael Stonebraker
Samuel Madden
Daniel J. Abadi
Stavros Harizopoulos
MIT CSAIL

{stonebraker, madden, dna,
stavros}@csail.mit.edu

Nabil Hachem
AvantGarde Consulting, LLC
nhachem@agdba.com

Pat Helland
Microsoft Corporation
phelland@microsoft.com

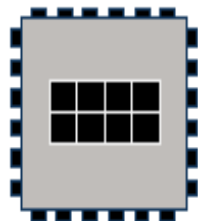
ABSTRACT

In previous papers [SC05, SBC+07], some of us predicted the end of “one size fits all” as a commercial relational DBMS paradigm. These papers presented reasons and experimental evidence that showed that the major RDBMS vendors can be outperformed by 1-2 orders of magnitude by specialized engines in the data warehouse, stream processing, text, and scientific database markets.

Assuming that specialized engines dominate these markets over time, the current relational DBMS code lines will be left with the

All three systems were architected more than 25 years ago, when hardware characteristics were much different than today. Processors are thousands of times faster and memories are thousands of times larger. Disk volumes have increased enormously, making it possible to keep essentially everything, if one chooses to. However, the bandwidth between disk and main memory has increased much more slowly. One would expect this relentless pace of technology to have changed the architecture of database systems dramatically over the last quarter of a century, but surprisingly the architecture of most DBMSs is essentially

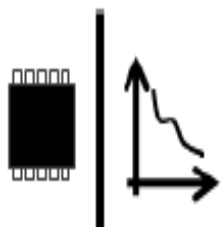
HW technology innovations



Multi-core architecture
(8 x 8 core CPU per blade)

Massive parallel scaling with many blades

One blade ~\$50,000
= 1 enterprise class server



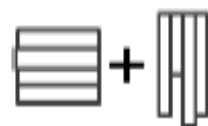
64 bit address space – 2 TB in
current servers

100 GB/s data throughput

Dramatic decline in
price/performance



SAP SW technology innovations



Row and column store



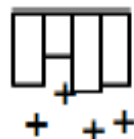
Compression



Partitioning



No aggregate tables



Insert only on delta

- Быстрая аналитика на больших многомерных данных, оперативность почти он-лайн. (3 часа – 16 сек, обновление 2-3 сек).
- Объем разработки и настройки меньше, чем для традиционных BI-решений. (не нужен ETL, нет агрегатов, витрин,...)
- Нагрузка на продуктивные системы снижена. (Младший в линейке серверов для HANA 512GB загружен на 10%CPU и 50%RAM.)
- Объем данных ниже в три раза, чем в BW.

Пилотное внедрение в Сургутнефтегазе - 2010 год, продуктив – 12.2011.

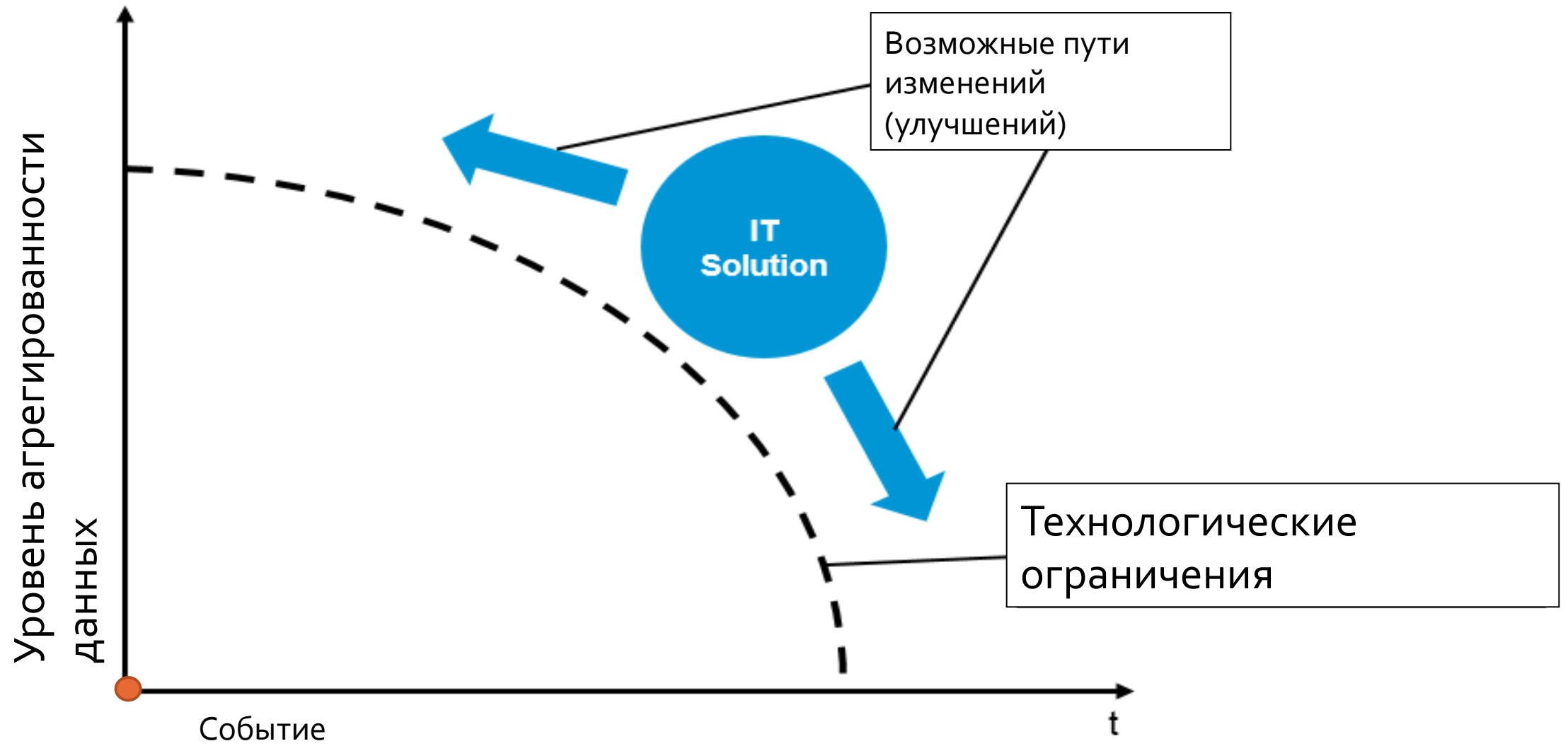
- Основной показатель улучшен на порядки.
- Дешевле
- Проще
- Смежные и синергетические эффекты.

IMDM – инновация, способная революционно изменить информационные системы.



Zero Latency
Real-Time Riding

Информация
полная, реакция
мгновенная.



The Real-Time Enterprise

"The RTE is an enterprise that competes by using up-to-date information to progressively remove delays to the management and execution of its critical business processes."

Progression is asymptotic — реальные предприятия всегда будут неэффективны по времени отклика. Оптимальный уровень задержки от реального времени — движущаяся цель — мы можем улучшать процессы по мере развития и удешевления технологий.

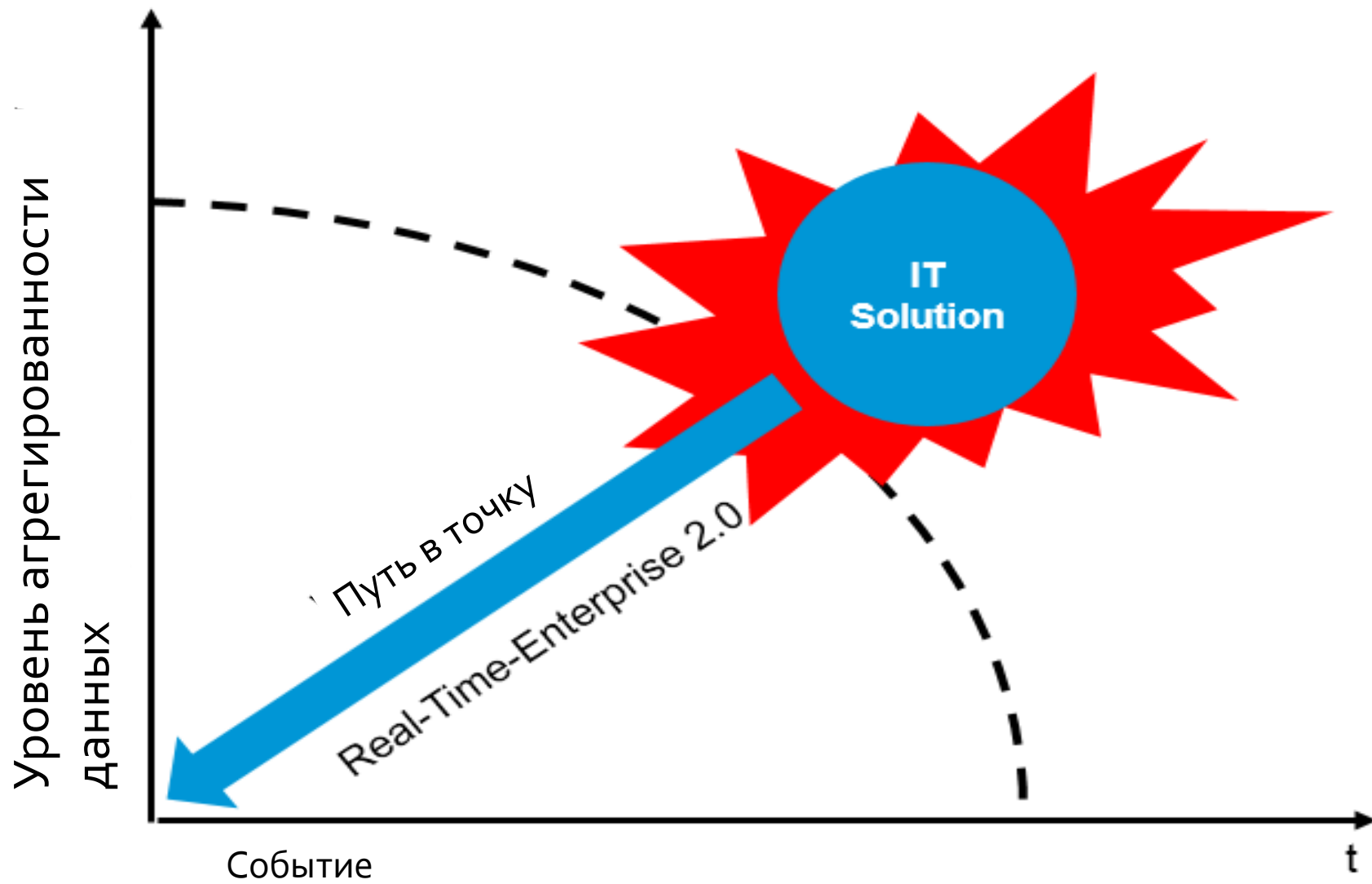
1 October 2002

COM-18-3057

Gartner.

The Real-Time Enterprise + IMDM: области изменений сегодня (2013-2018)

- Одна БД (СУБД). Конвергенция ОЛТП и ОЛАП.
- Подходы к программированию.
- Модели данных, модели процессов.
- Безопасность информации.
- Инфраструктура. ИТ-архитектура.
- Обработка потока событий.
- СППР + управление жизненным циклом решений.
- Бизнес-процессы и бизнес-системы (финансы, логистика, производство, HR, ...).
- Бизнес-модели.
- Enterprise Architecture (в совокупности и по доменам).



The Real-Time Enterprise 2.0

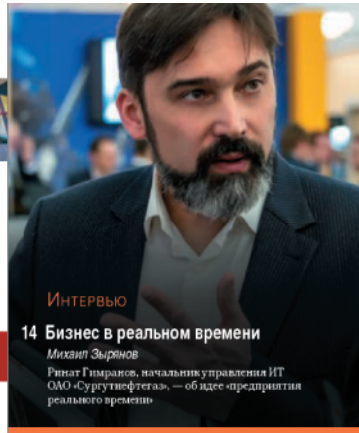
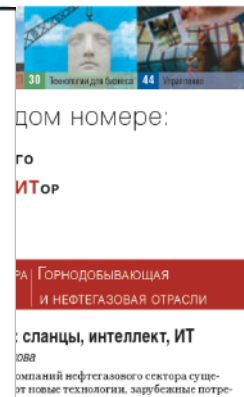
1.0: "The RTE is an enterprise that competes by using up-to-date information to progressively remove delays to the management and execution of its critical business processes."

2.0: "The RTE is an enterprise that competes by using up-to-date information to completely remove delays to the management and execution of its business processes."

- О наших проектах на SAP HANA см. сайты SAP, Нефтяное хозяйство,...



05 Май 2013



ПОДДЕРЖКА

«Сдвиг парадигмы»: IMDM существенно изменит корпоративные информационные системы, окажет (оказывает!) сильное влияние на ИТ-отрасль и рынок.

Компетенции: технология.

• Специалисты по БД (для IM-DBMS):

- Знать технологию: особенности, подходы, принципы.
- Знать текущие программные решения (SAP HANA, Oracle EXADData, IBM SolidDB,...)
- уметь проектировать и оптимизировать структуры данных.
- Уметь разрабатывать и отлаживать новые SQL-запросы.

• Специалисты по BI (Business Intelligence):

- Знать технологию.
- Знать особенности построения решений IM.
- Уметь проектировать и оптимизировать юниверсы.
- Уметь выбирать и проектировать аналитические решения (отчетность, СППР, обр-ка событий, симуляторы, ...)

• Специалисты по интеграции:

- Знать технологию.
- Знать особенности обеспечения интеграции на уровне систем, процессов, событий, данных.
- Уметь проектировать и оптимизировать интеграционные решения в составе ИТ-архитектуры.

• Специалисты по инфраструктуре:

- Знать технологию.
- Знать варианты построения ландшафта систем с использованием решений IM – компоненты, связи, существенные параметры.
- Уметь проектировать и оптимизировать ландшафт.
- Уметь выполнять сайзинг.
- Уметь разрабатывать стратегию развития ландшафта.

Компетенции: бизнес.

- Специалисты по БП:
 - Знать ...
 - Уметь...
- Специалисты по ИТ-архитектуре:
 - Знать...
 - Уметь...
- Специалисты по Enterprise Architecture:
 - Знать...
 - Уметь...
- Специалисты по ?



Пространство для научных и практических работ.

МГУ

win-win и
синергия.
1 + 1 + 1 = 5!

СурГУ

Сургут-
нефте-
газ

Базовая кафедра ОАО «Сургутнефтегаз»:

• ВУЗ:

- Повышение качества подготовки бакалавров
- И магистров
- Инновационные, интересные и полезные научные темы
- Повышение квалификации ППС.
- «Чувство» современного производства – прямую и обратную связь.

• Предприятие:

- Повышение наукоемкости (= эффективности) проектов по развитию информационной системы.
- Повышение квалификации в рамках системы непрерывного образования (практикант – магистр – аспирант – преподаватель/эксперт).
- Дополнительная мотивация высококвалифицированного персонала.

ИТ - есть понятные требования к компетенциям специалистов. Нет (пока) осуществления.
Бизнес-задачи - только потенциальные возможности (challenges).

Спасибо за
внимание!

Gimranov_rd@Surgutneftegas.ru





«Сегодня, в непростых экономических условиях, когда мы должны ускоренными темпами развивать несырьевые отрасли, на первый план, конечно, выходит максимальное сближение образования с производством, с практикой, с реальной жизнью. Создание прорывных технологий, обеспечение производств востребованными кадрами – это наша общая задача, и, конечно, такая связь с производством просто необходима.»

В.В.Путин

«Главный вопрос нашей повестки дня: каким путём идти дальше нашей высшей школе? Требуются ли поправки к некоторым траекториям нашего движения? Что мешает нам более эффективно работать? И, главное, какую ответственность мы можем и должны взять на себя?»

«К сожалению, разорвана важная цепь: образование – фундаментальные исследования – научный результат – внедрение его в технологию и производство. Отраслевые институты исчезли, а замены им мы пока не нашли. Хорошей альтернативой могут стать университетские научно-технологические долины, кластеры, технопарки, тесная связь с корпорациями, но, к сожалению, это ещё не вошло в нашу жизнь и не приносит, на наш взгляд, нужных результатов.»

В.А.Садовничий