IT Growth Beyond Commodity-Processor Acceleration

Dr. Karl-Heinz Strassemeyer

Distinguished Engineer System Platforms
Member of the IBM Academy of Technology
IBM Systems & Technology Group, Boeblingen
strasse@de.ibm.com

Intelligence Beats Speed

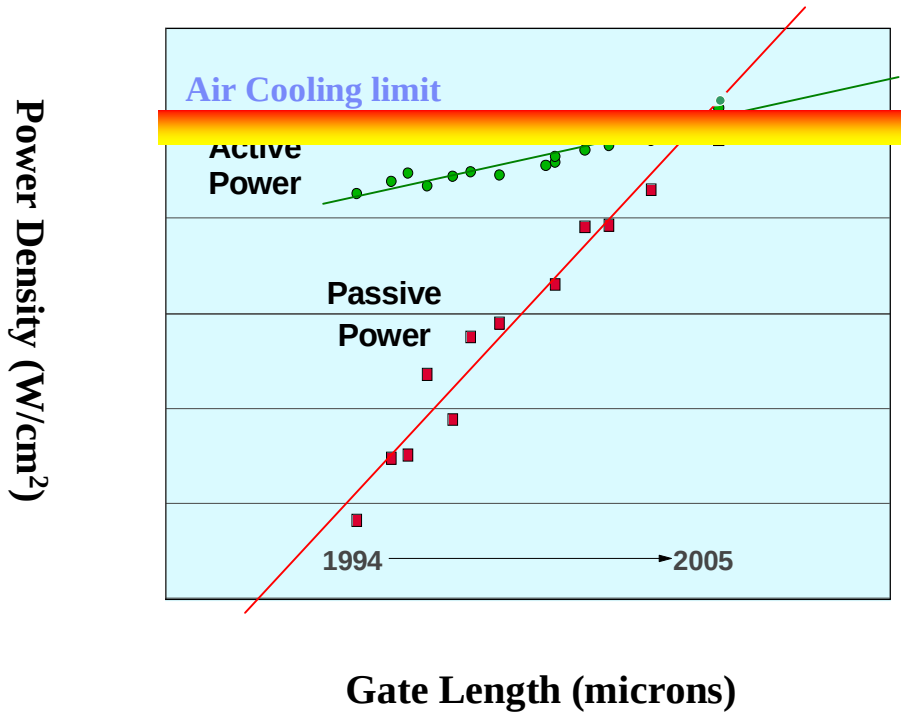
Processors
Penetrating the Technology Wall

Intelligence Beats Speed

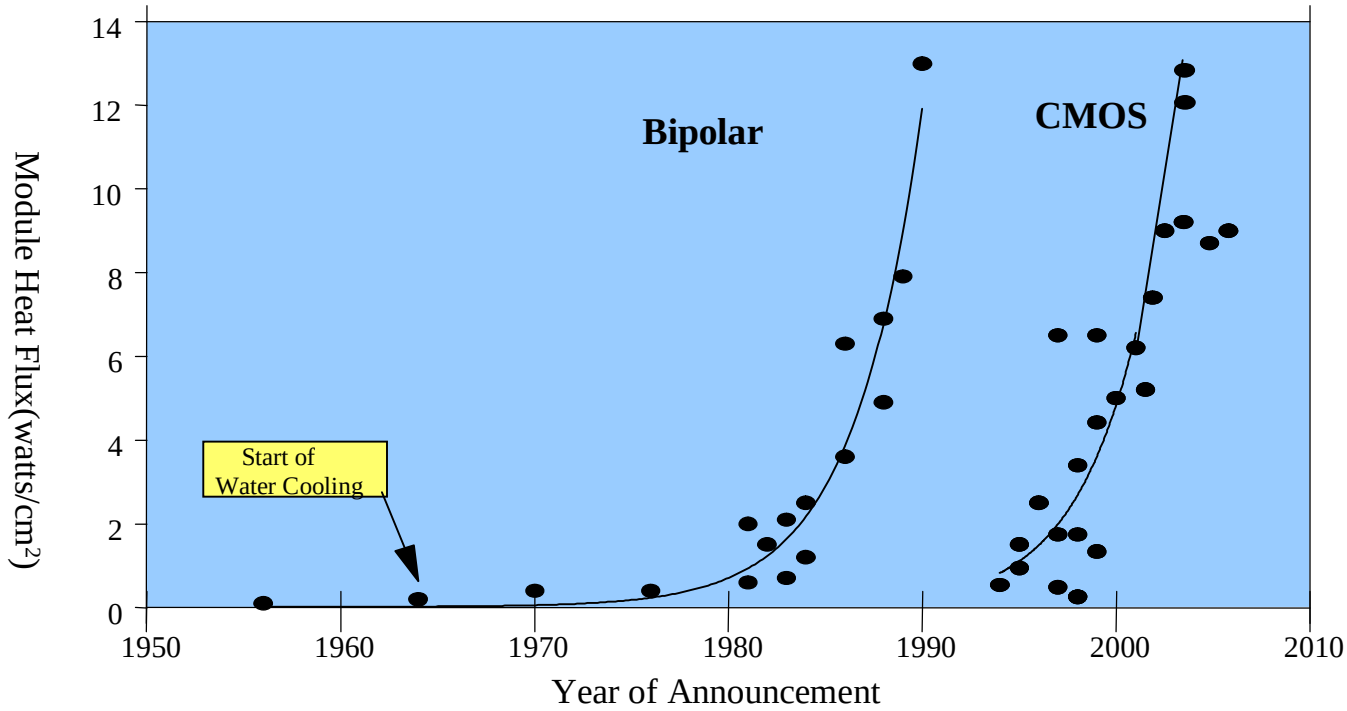
CMOS Power Issue: Active vs. Passive Power

Power components:

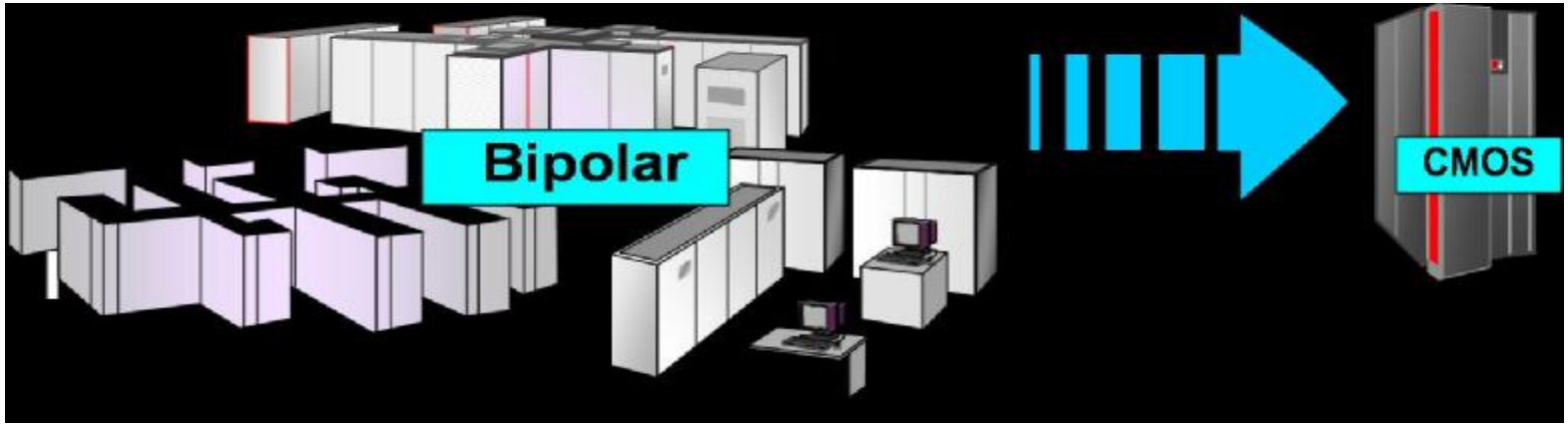
- Active power
- Passive power
  - Gate leakage
  - Source – Drain sub-Vt leakage



Technology Discontinuity: CMOS Power Crisis

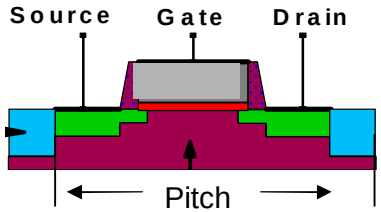


Bipolar to CMOS



Maintenance:	**up to 65% reduced**
Energy:	**up to 97% reduced**
Area:	**up to 91% reduced**

Possible Technology Roadmap

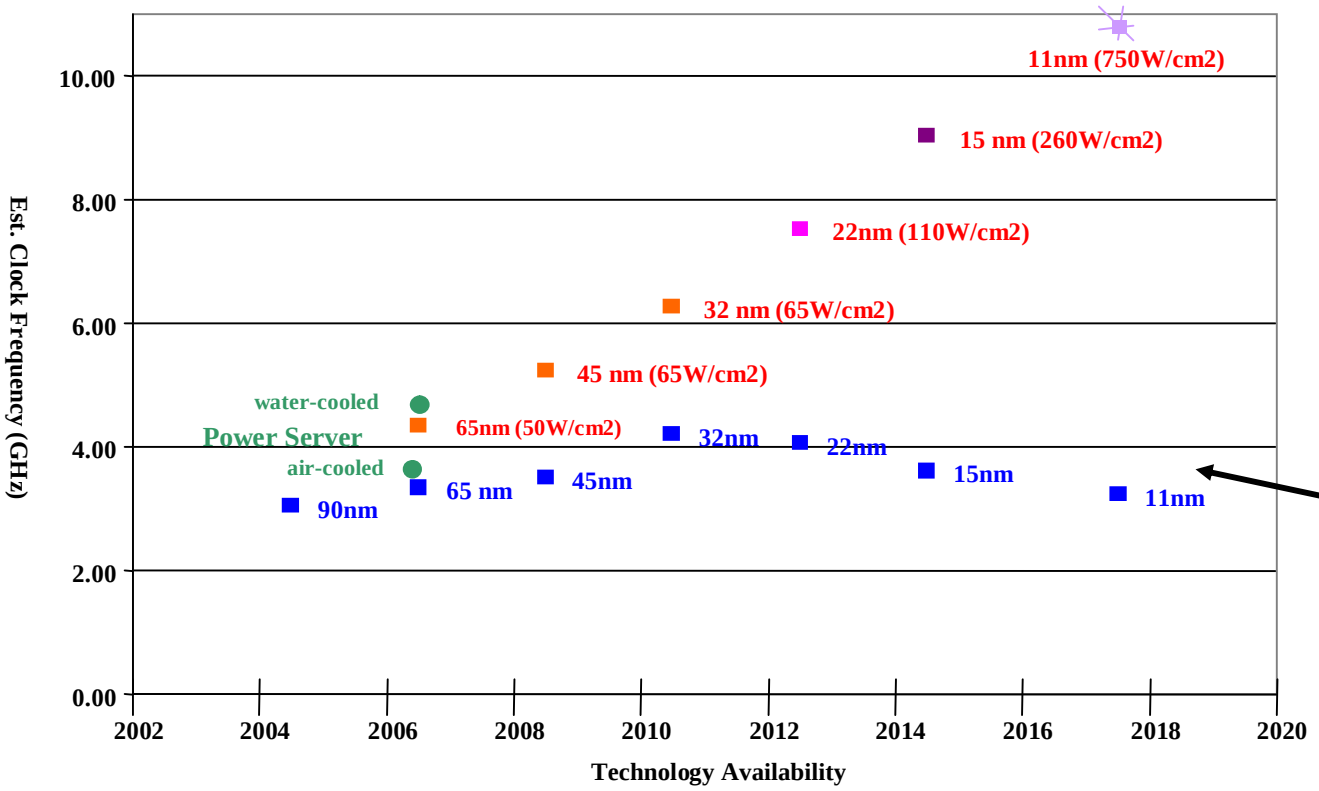


§ **Assuming lithography is capable** UV 193 nm EUV 43 nm?

Node	Device Pitch (nm)	Minimum W (nm)	Nominal L (nm)	SRAM Cell Size (μm^2)	Technology Availab. Year
45	170-180	70-80	40-45		2008
32	120-130	50-60	30-35		2010
22	80-100	35-45	25-30	0.08-0.1	2012
15	65-75	25-35	20-23	0.04-0.06	2014-15
11	45-55	16-25	12-18	0.02-0.04	2017-18
8	35-45	10-16	9-16		
5	25-35	7-10	7-12		

Frequency Scaling to 11 nm

Optimizing for maximum performance for each core

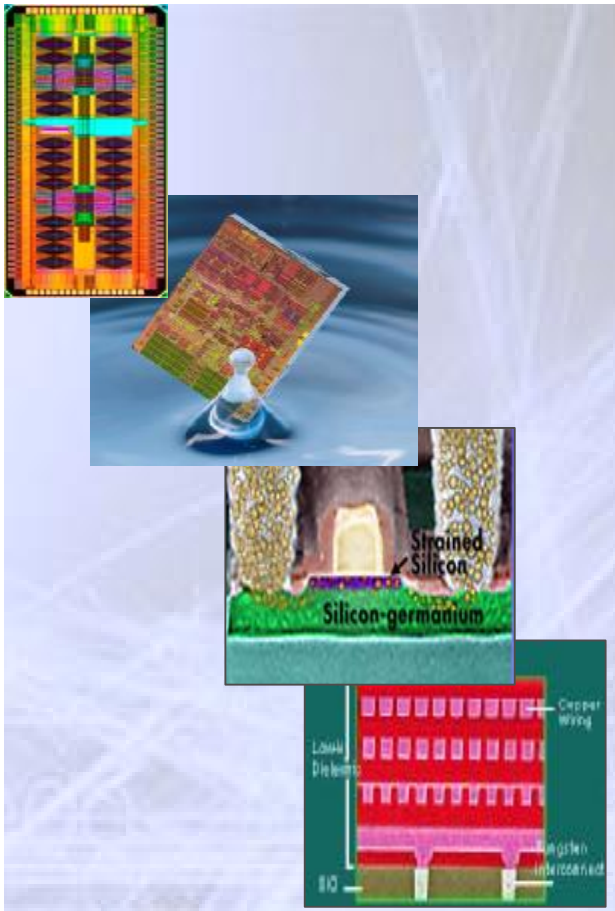


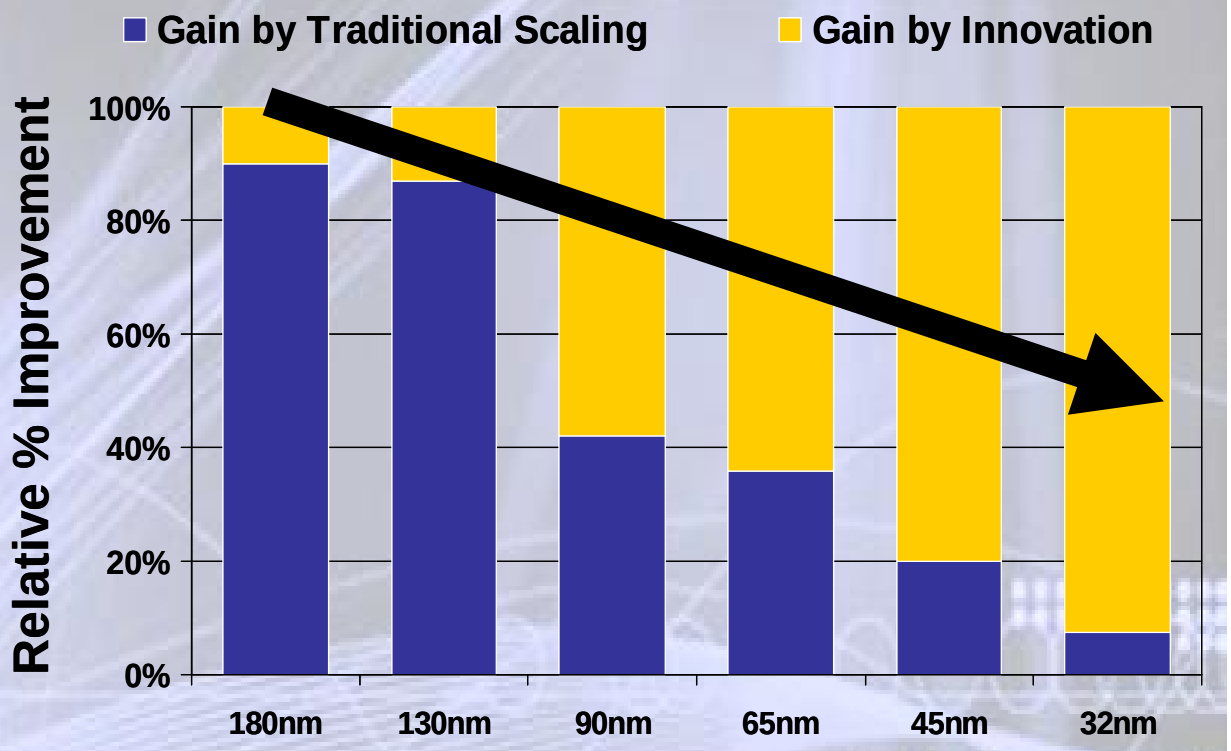
Constant performance improvement, 20% per gen.

Constant power density 25 W/cm^2

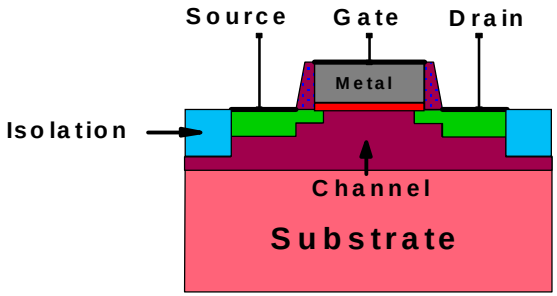
D. Frank, C. Tyberg

Technology Scaling… The Fuel of Exponential Growth





High-K & Metal Gate



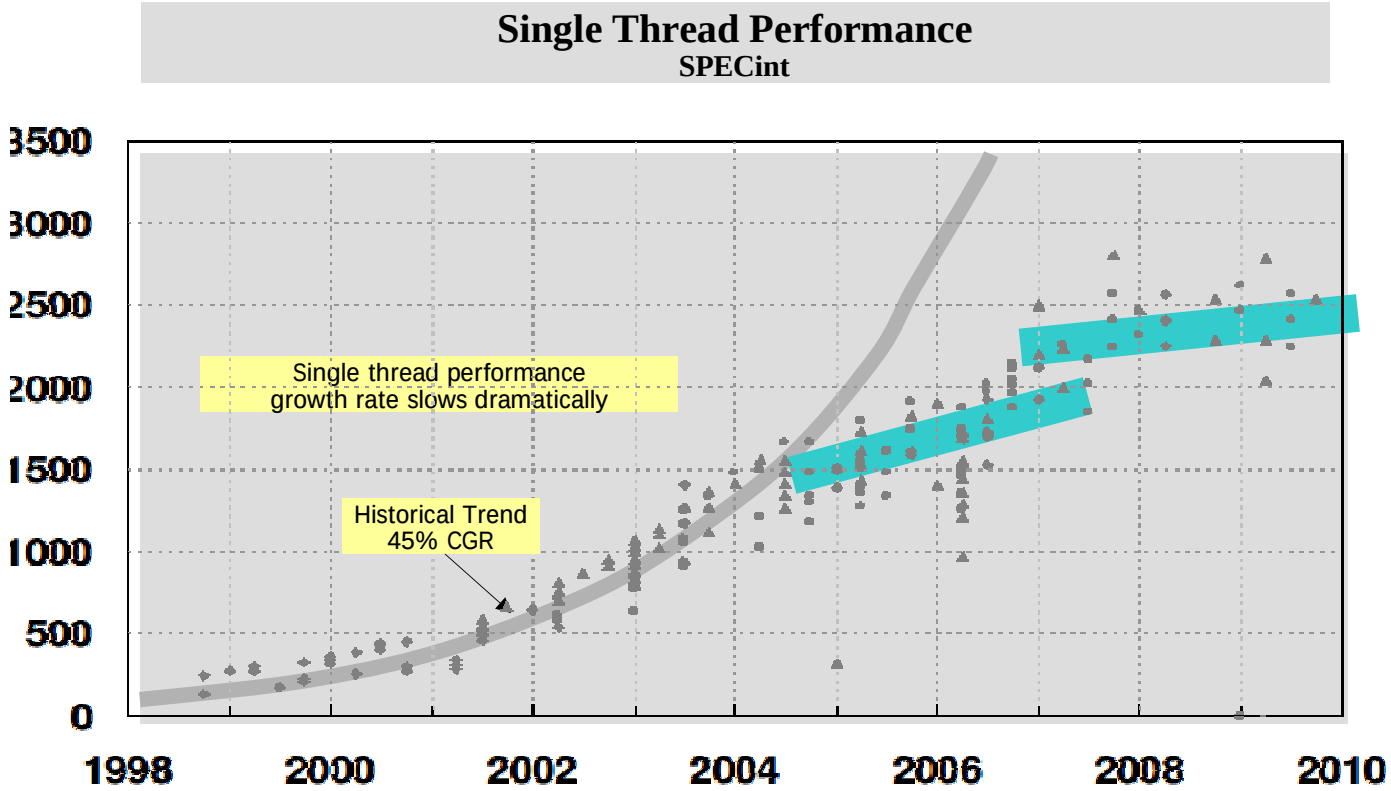
45 nm CMOS Process

⇒ Hafnium-based gate dielectric (High-K) + Metal Gate

⇒ Same capacitance with thicker dielectric

⇒ Gate leakage reduction and higher performance
or same performance with much lower total power

System on a Chip – Multicore Chips



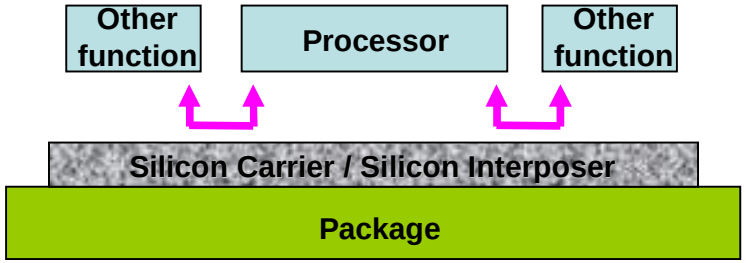
Other Technology Innovations

Intelligence Beats Speed

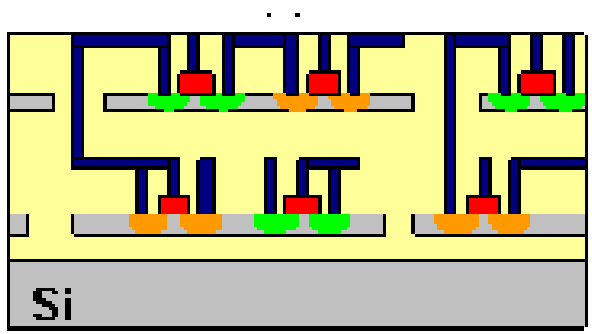
Exploit the 3rd Dimension

§ A family of technologies enabling stacking of active Si layers and vertical connections between them.

Silicon Carriers / Interposers



3D Integration



Die assembly
Low via density
(< 10^5 cm^{-2})

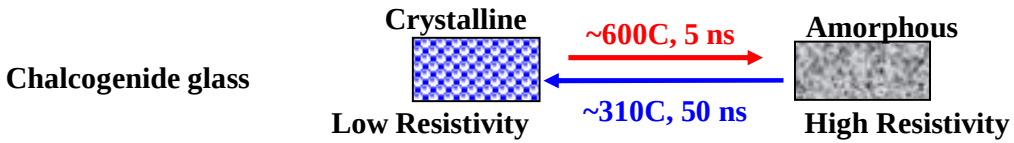
Time

Wafer assembly
High via density
(up to ~ 10^8 cm^{-2})

Integrated DRAM Cache

- Large cache improves system performance
- On-Chip cache eliminates I/O delay
- Multi-core designs further stress caches
- DRAM offers 3-4x more bits / mm^2 than SRAM
- Slower but smaller = less wire delay
- DRAM keep alive power ~ 1/5 of SRAM
- DRAM Soft Error Rate ~ 1/1000 to 1/5000 of SRAM
- Implemented with 45 nm Technology

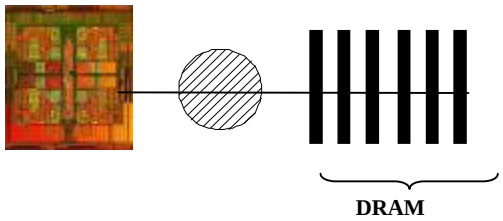
Phase Change Memory (PCM)



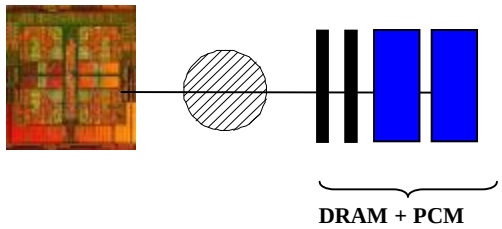
- Material can be heated to change between amorphous and crystalline states
- Read Latency: ~100ns; Write Latency: ~1000ns
- 100 times density of DRAM

Replace Flash in solid State Disks
Mixed Memory Architectures

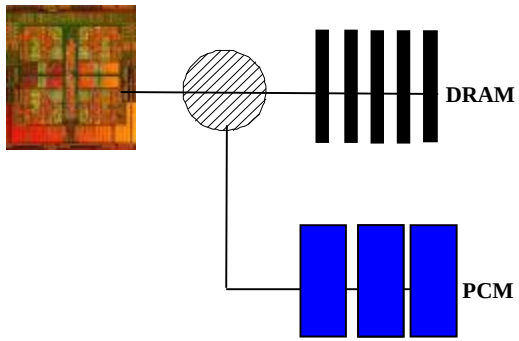
All DRAM Memory System



Mixed Memory System



Dual Memory Architecture



Future Solution Focus

Intelligence Beats Speed

Mike Rhodin - Vice President IBM SW Solutions

Boeblingen Lab Visit July 14th, 2009 – at that time General Manager Northeast Europe

IBM does since hundred years the same thing:
Help customers to solve their important problems with technology

We started with the Chicago Meat Industry

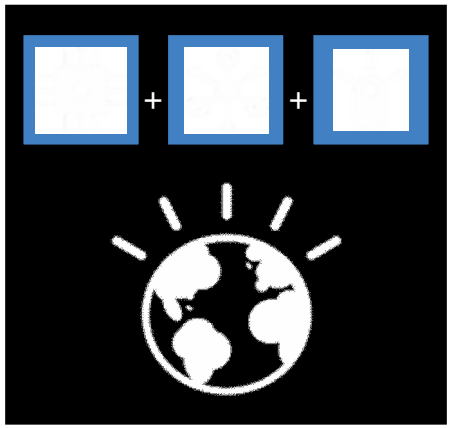
In the fifties Thomas J. Watson focused us on the theme:
Automation of the Banking Industry

Last year Sam Palmisano set the goal for the next fifty years:
Creation of a Smarter Planet

Evolution of the Smarter Planet

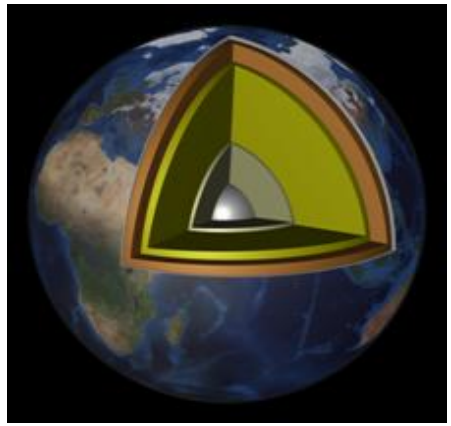
Smarter Planet (Past)

Distinct physical, people, IT and business worlds



Smarter Planet (Present)

Instrumented, interconnected, and intelligent

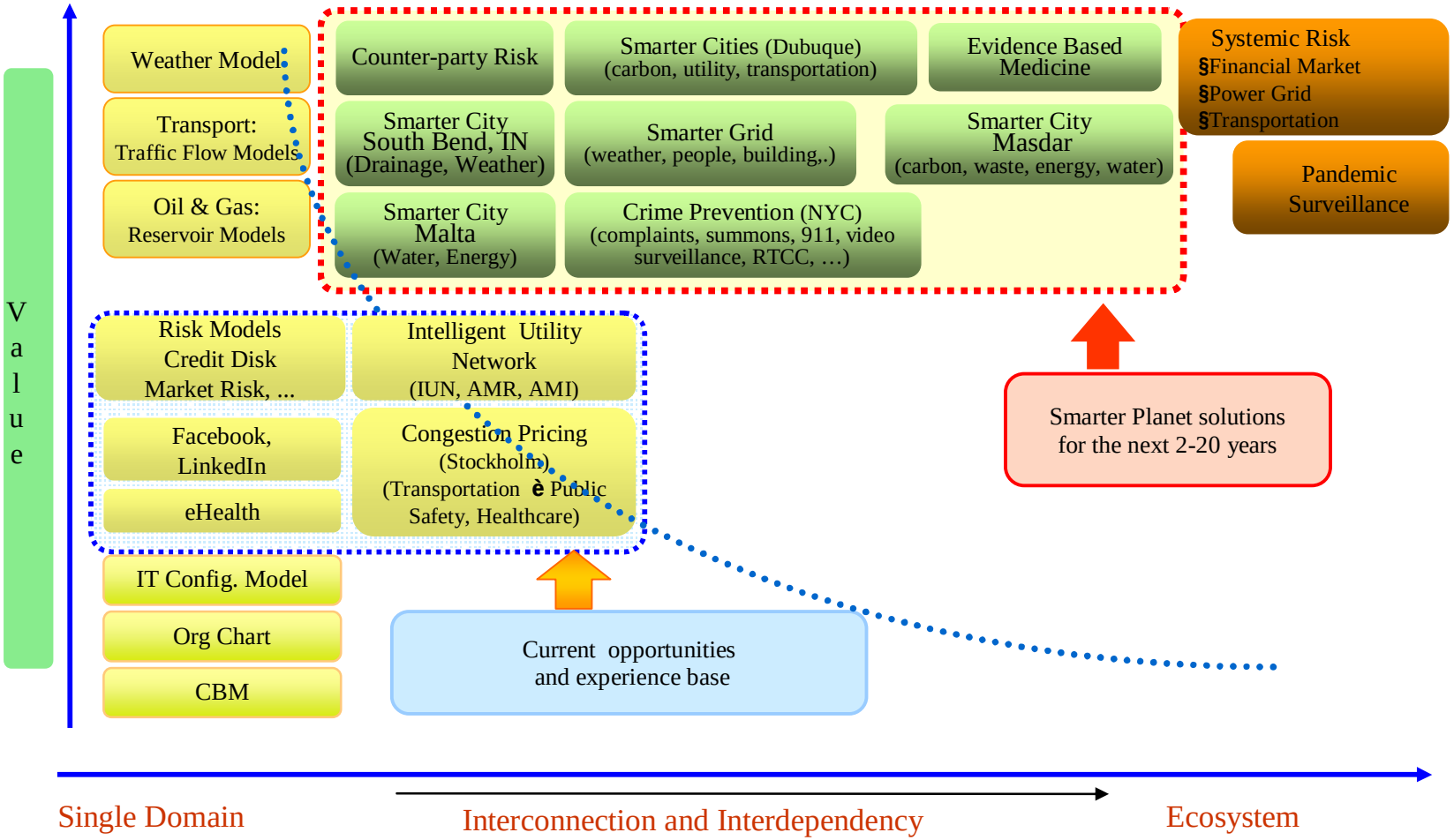


Smarter Planet (Future)

Interactive, interconnected & interdependent, digitally represented world

- **Interconnected and interdependent behavioral models optimize Smarter Planet solutions**
- **Dynamic capture and assimilation of data using closed-loop models for prediction & response**
- **Individual and community behaviors & preferences leveraged for improved business outcomes**

Smarter Planet solutions



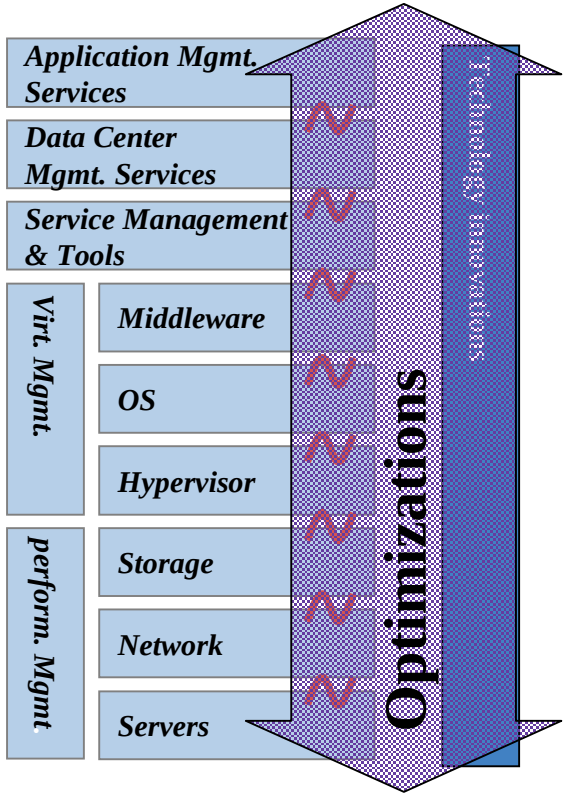
Workload Optimized Systems

Intelligence Beats Speed

Workload Optimized Systems

A Workload Optimized System:

- **provides value for particular workload or set of workloads important to the client**
- **provides unique functionality or differentiated performance**
- **reduces cost of deployment and operation**
- **is accomplished through co-design of HW, SW and services**



Workload Optimized Systems

Workload Optimized Systems deliver transformational improvements in client value (time to value, operating costs, scale/performance) for important application domains.

These workload optimized systems will require system-wide co-optimization of hardware and software. They exhibit a tight connection between software, compute elements, extended memory and storage elements, and high-speed network.

Successive refinement and expansion of customer value over a multi-year roadmap is required to overcome the inertia of current industry practices of component level standardization.

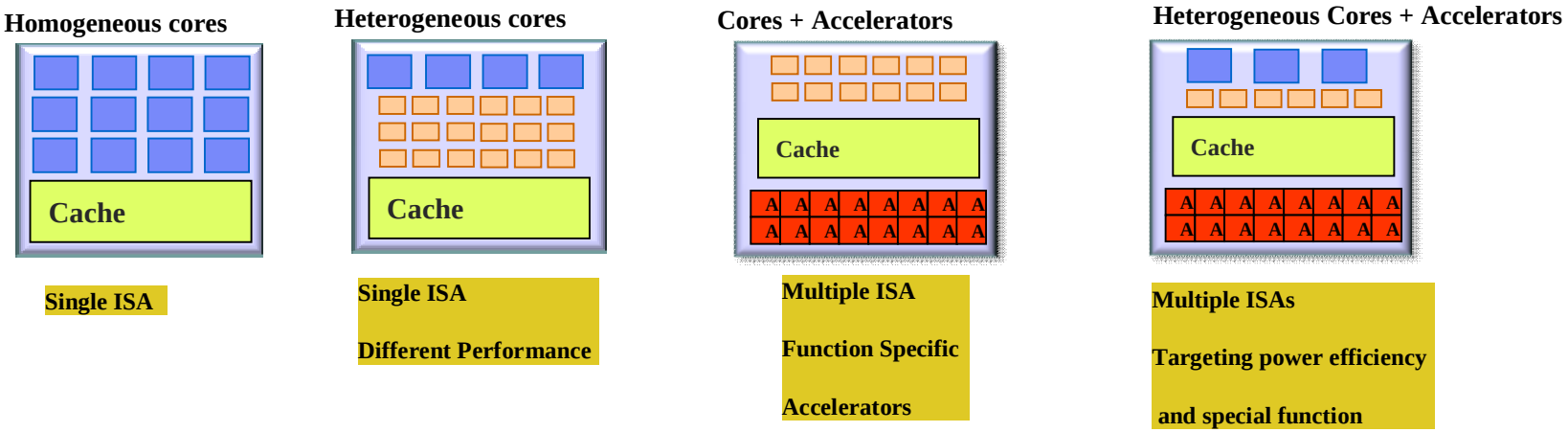
A small set of Workload Optimized Platforms will emerge to support important classes of Workload Optimized Systems.

Chip Level Architecture

Many options in chip-level architecture will be available:

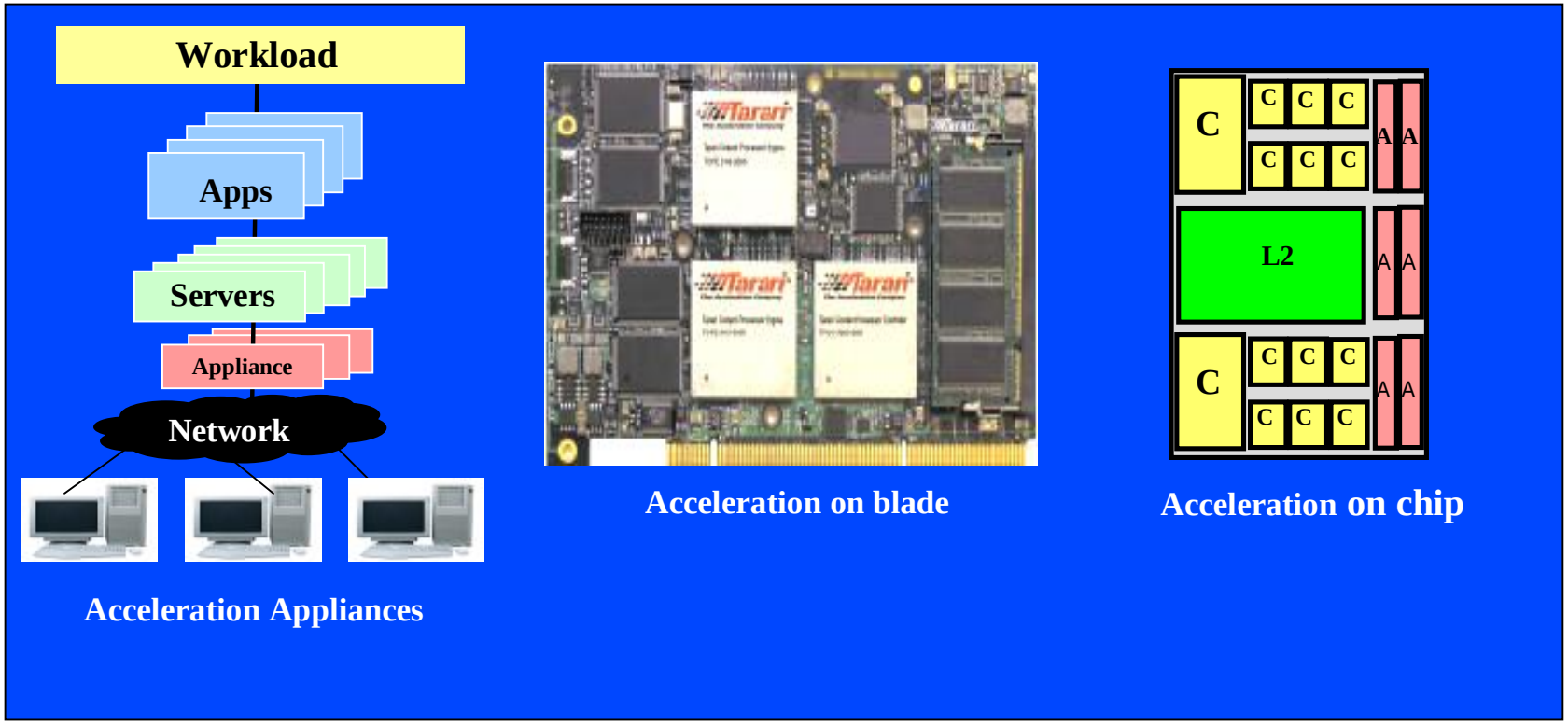
- Number and types of cores
- Memory hierarchy
- Interconnect structure

- **Optimization for Power vs Performance will be important**
- **Accelerators and heterogeneity will be exploited to optimize for workload specific special functions**



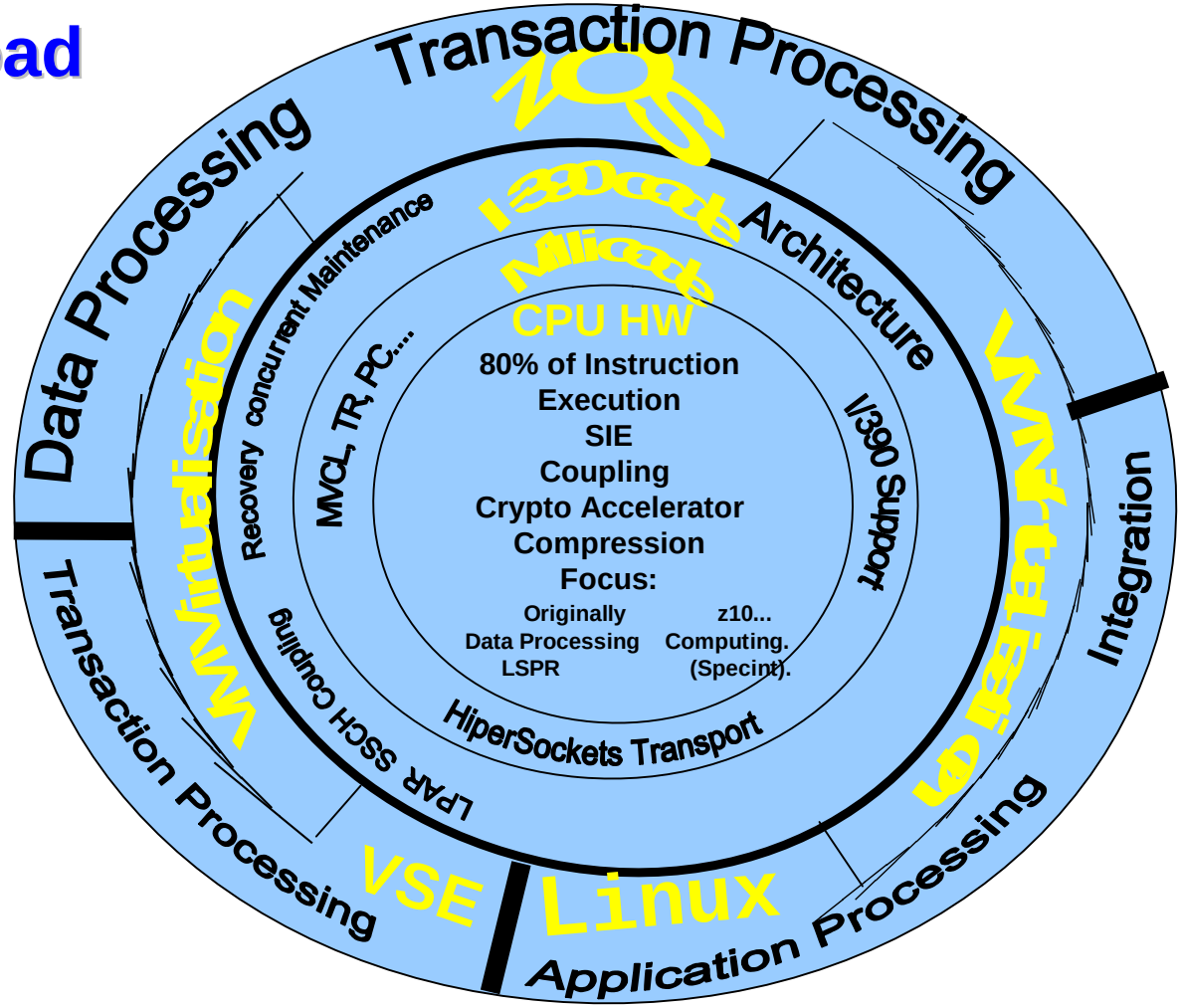
Accelerators

- Hardware / software units, customized and integrated to deliver client value in a specific solution area
- Implemented at all Levels

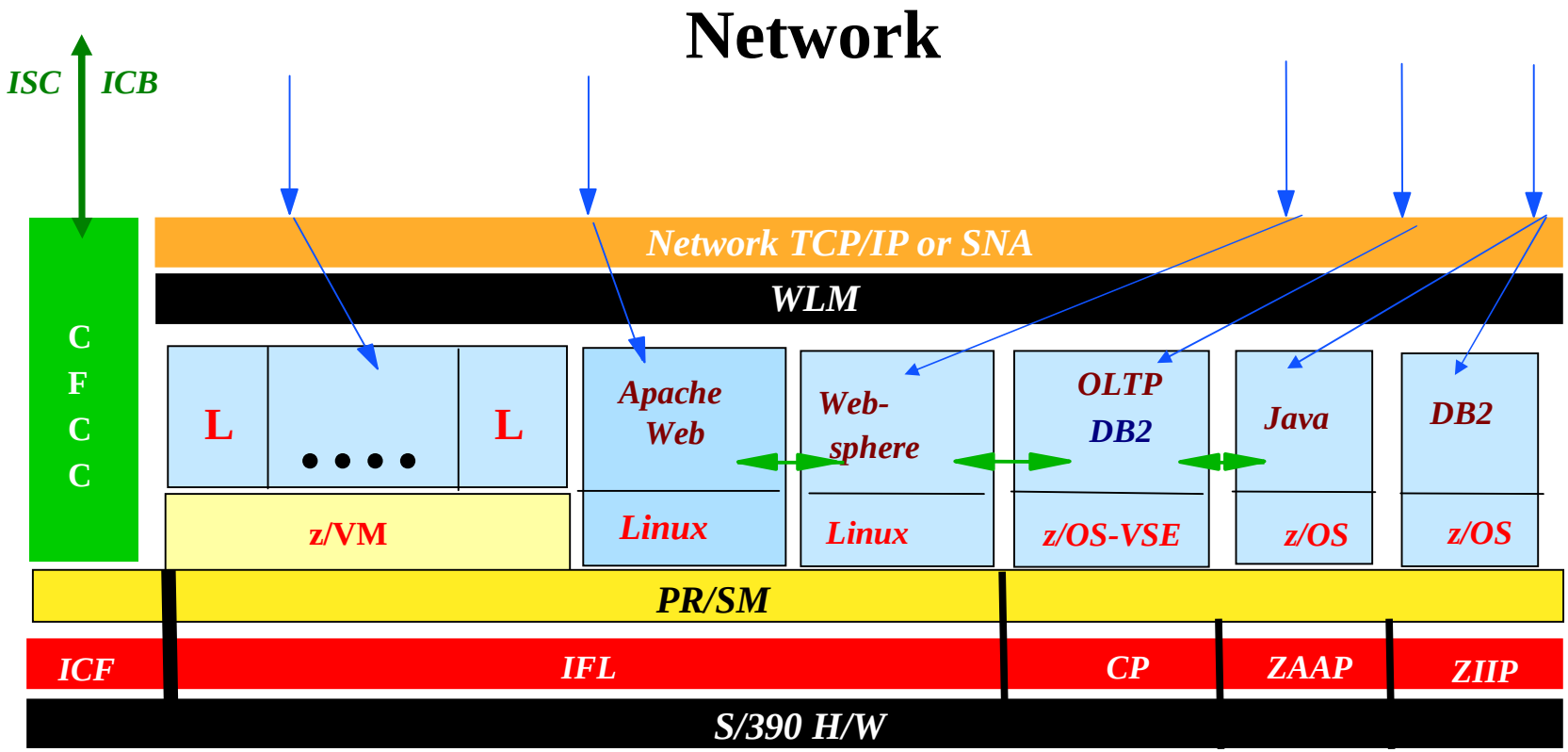


System z, a Workload Optimized System

since 45 years



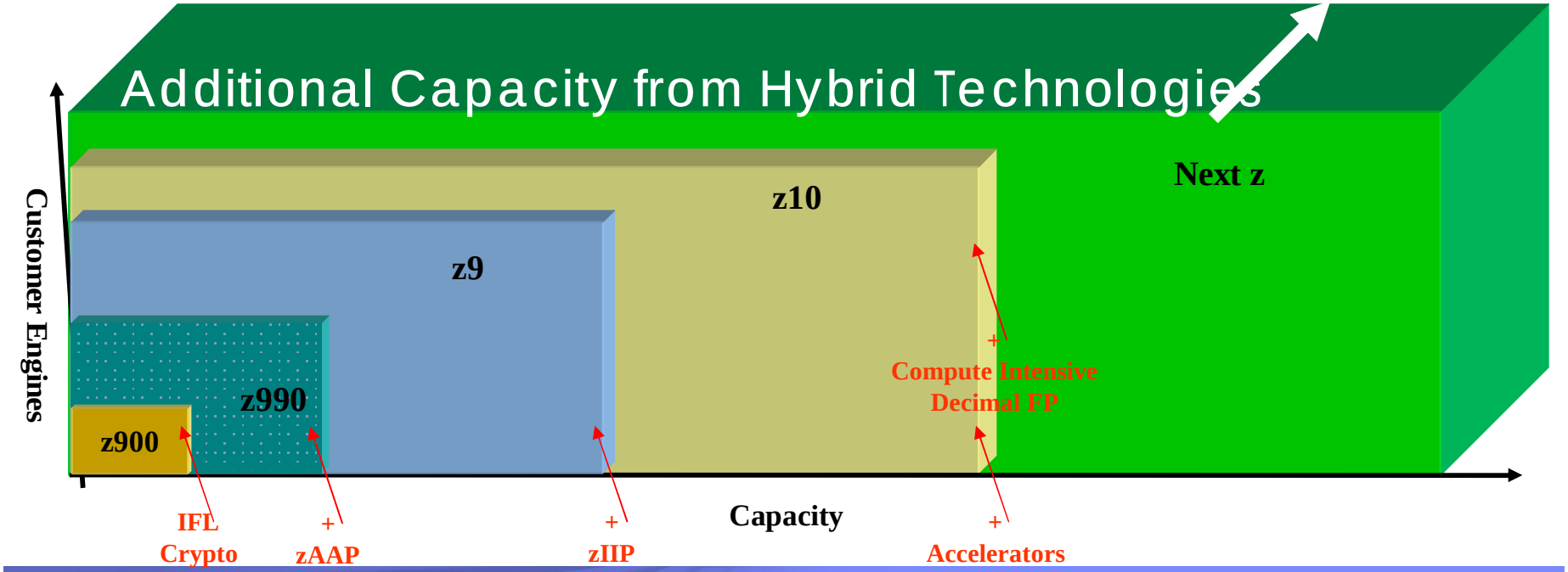
zSeries Platform Exploitation



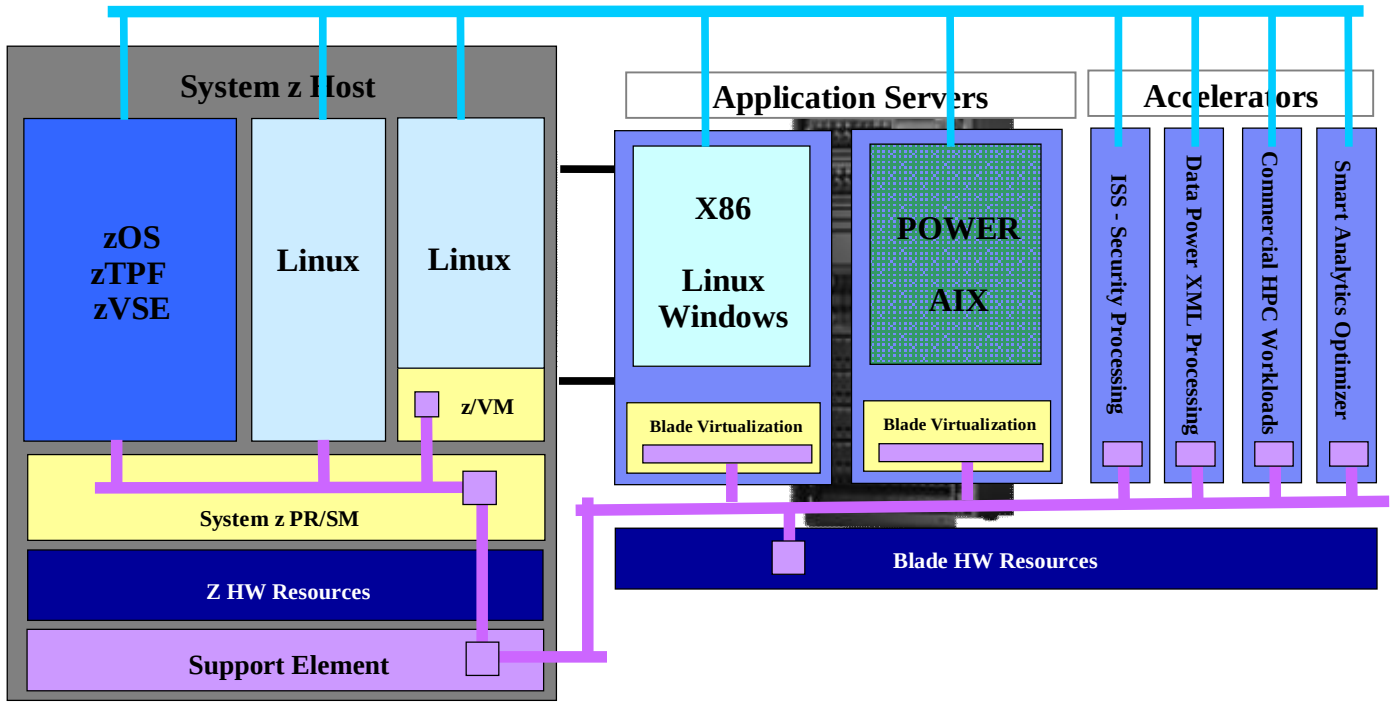
Processor-Performance and Scalability

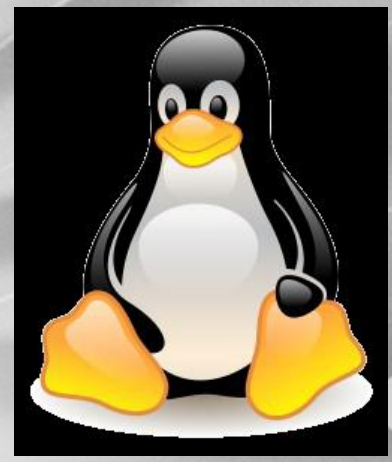
Ø extend the scalability beyond the traditional growth

Ø zFuture will take advantage of extended integrated Technologies



Expand Systems Management and Connectivity





Ten Years Linux on the Mainframe

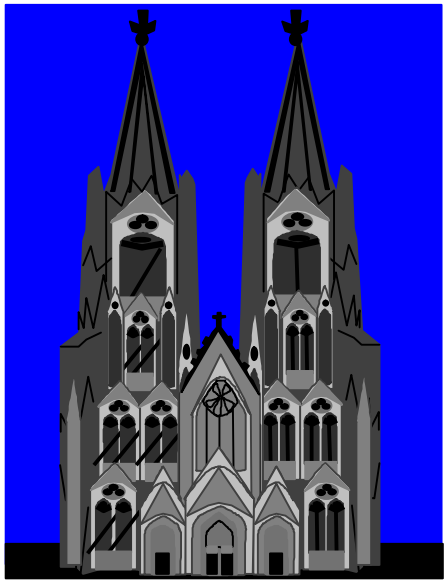
The Beginning of a Heterogeneous Data Center in one System

Intelligence Beats Speed

Do it the LINUX way

Established development Process:

â Cathedral Style



n A different Culture: Bazaar flexible (re-) organization

è dynamic processes
è contents always up-to-date
è all tasks in parallel
è no idling
è designed by participants

From Skunk-Work in Boeblingen to a Strategic Product

UNIX X on S/390: AIX Unix System Services on OS/390 Auto Unix

May 1998: Started Linux study VSE Team GCC Compiler available (microcode)
Oct. 12th 1998: IBM Academy of Technology Birds of Feather session (Java just in time compiler)
Jan. 1999: Feasibility Established (Kernel + 5 I/O drivers)

Focus on baby /390 Target Sept. 1999

Aug. 16th 1999: Letter to "all IBM executives": Linus Torvalds body language indicated: "Linux on Systems /390 no good idea"
Two weeks later: Meeting with Linus Torvalds Santa Clara Marriott: He was enthusiastic Body language: no business talk in a disco

Focus on IBM top down strategy

Oct. 1999 WAVV Meeting
Early Nov. 1999: "Show" code to Linus Torvalds promised Release into OpenSource 1999
Dec. 15th 1999: Release into developerWorks
Dec. 16th 1999: Linux 1.16.1 including Patch for Mainframes from IBM
Jan. 2000: Linux World New York Palmisano 1B$ into the ECO System
Linus Torvalds 1B$ is a lot of money but not that mutch
May 15th 20001: Announcement in Palm Springs

Design Principles

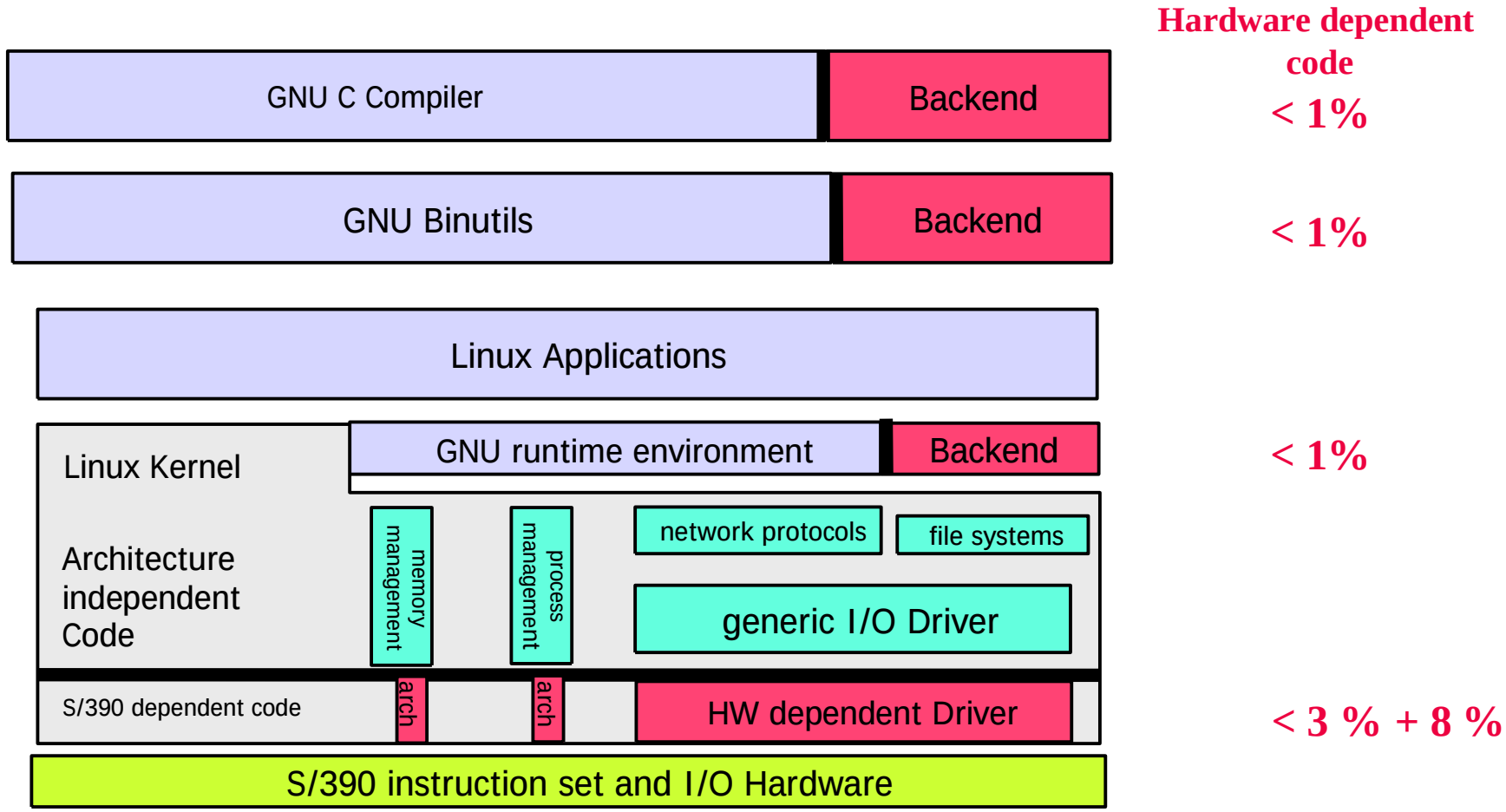
Ø Linux/390 remains Linux

- The Linux structure, development rules and coding style remain unchanged

Ø S/390 remains S/390

- The mere S/390 hardware architecture is sufficient for implementing Linux

Linux – the Plattform for Application Development and Execution



Middleware Support 2001

2.4 Kernel/glibc 2.2, 31 bit

DB2 UDB (incl. DB2 Connect)
CICS Transaction Gateway
IMS Connect
MQ Client (C) / Server (S)
Java JDK
WebSphere
- Commerce Suite Pro Edition
- Portal Enable Solution
- Edge Server

Tivoli TSM Client (C) / Server (S)
Tivoli Policy Director
Lotus Domino Enterprise Server

4Q01	1Q02	2Q02	3Q02	4Q02	1Q03
V7.2		V8			
		V4			
V1.1					
V5.2 C				V5.3 S	
V1.3.0	V1.3.1				
		V4.0.3	V5		
			V5.2		
		V4			
		V4.2 C			V4.2 S
		V4.0			
			V6		

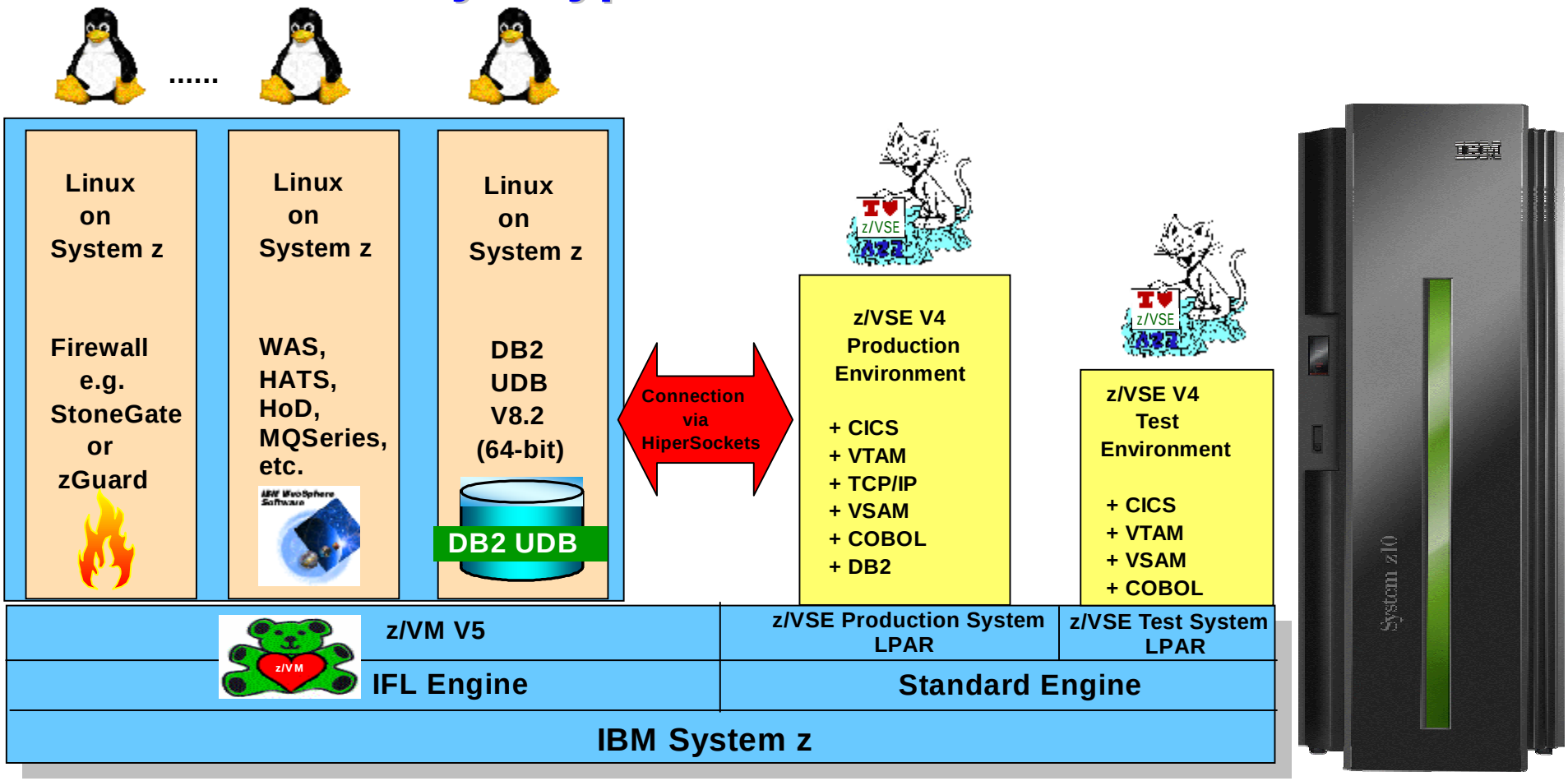
Colour coding: available in plan under consideration

Middleware and Applications on Mainframe Linux now

- **Information management software**
  - Cognos Business intelligence
  - Content Manager
  - DB2 9 for Linux, UNIX and Windows
  - DB2 Connect, features and tools
  - DB2 Enterprise Server Edition
  - FileNet product family
  - Informix Dynamic Server
- **WebSphere software**
  - WebSphere Application Server
  - WebSphere Business Monitor
  - WebSphere Commerce
  - WebSphere Dynamic Process Edition
  - WebSphere Enterprise Service Bus
  - WebSphere Event Broker, WebSphere Message Broker
  - WebSphere Federation Server
  - WebSphere Process Server
  - WebSphere Process Choreographer
  - WebSphere Adapters
  - WebSphere MQ
  - Communication Server
- **Lotus software**
  - Lotus Domino
  - Lotus Connections
  - WebSphere Portal / Portlet Factory
- **Rational software**
  - Rational Developer for System z
  - Rational Host Access Transformation Services
  - Rational Quality Manager
  - Rational Team Concert
  - Rational Asset Manager
  - Rational ClearQuest Mulitplatform
- **Tivoli software**
  - Maximo Asset Manager
  - Tivoli Access Manager
  - Tivoli Business Service Manager
  - Tivoli Common Inventory Technology
  - Tivoli Configuration Manager
  - Tivoli Enterprise Console
  - Tivoli Monitoring
  - Tivoli Netcool
  - Tivoli Omegamon XE
  - Tivoli Provisoning Manager
  - Tivoli Risk Manager
  - Tivoli Storage Manager
  - Tivoli System Automation
  - Tivoli Workload Scheduler

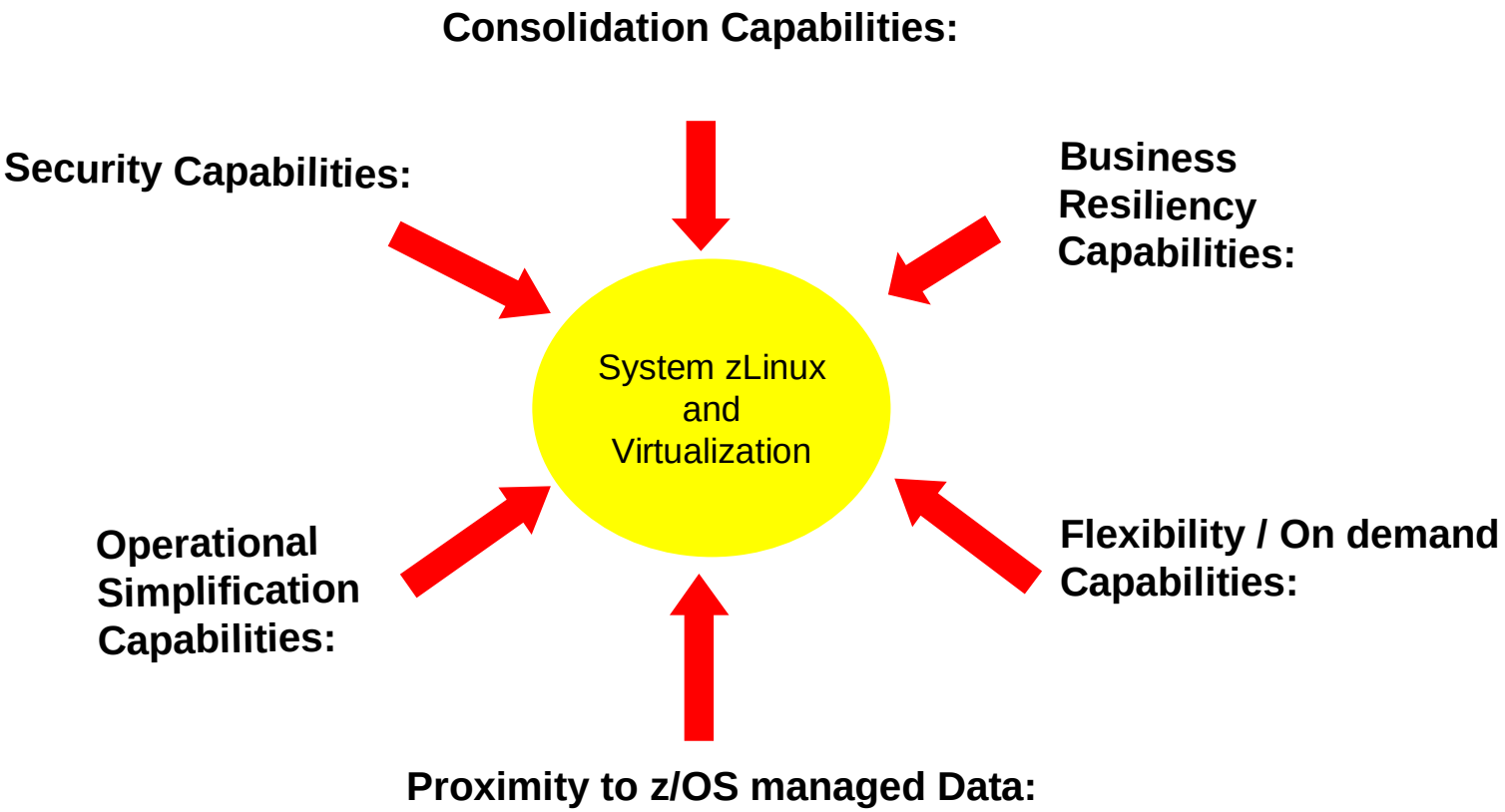
See the IBM Middleware Matrix for more products:
http://www.ibm.com/linux/matrix/

Today‘s typical z/VSE Environment



Linux is Linux… *but…*

System z provides unmatched value propositions to Linux workloads



Integration with Linux – Customer Examples

The Home Depot

Second largest retailer in the United States, over 1800 stores, 300,000 employees and around 1.2 billion customer transactions a year. Revenue 90Billion

Solution / Benefits:

- **Migrated SAP R/3 to DB2 on z/OS**
- **Near continuous operations**
- **Strategic investment to move retail apps to SAP retail**
- **Moving application servers to Linux on z for provisioning ease**
- **Faster time to market for new offerings in stores**
- **SAP BW with operational data**

"For The Home Depot, given our size and our requirements, IBM System z is the only choice."
Jim Fisher, Home Depot

Nationwide

Fortune 100 insurance & financial services company

Solution:

- **Consolidated Intel and UNIX application servers to Linux on System z - 478 virtual Linux servers**
- **12 mission critical applications deployed to Linux on System z - 100,000+ active users every day**

Benefits

- **Better TCO ($15 million savings over 3 years)**
  - 50% reduction in Web hosting monthly costs
  - 80% reduction in data center floor space needs; power conservation
  - 50% reduction in hardware & OS support efforts
  - Significant savings on middleware costs; WebSphere, DB2 UDB, Oracle
- **Faster provisioning speed (months to days)**
  - Dynamic allocation of compute power
  - Capacity on demand; increase/reduce compute power
  - Simple and robust high availability & disaster recovery

"Nationwide's Linux on System z project is currently estimated to save $16 million dollars over the next three years, not including floor space. We also were able to provide a reduction in server cost of more than 50 percent to our customers. The Linux on System z system saved significant data center floor space and power consumption."
Steve Womer, Senior IT Architect

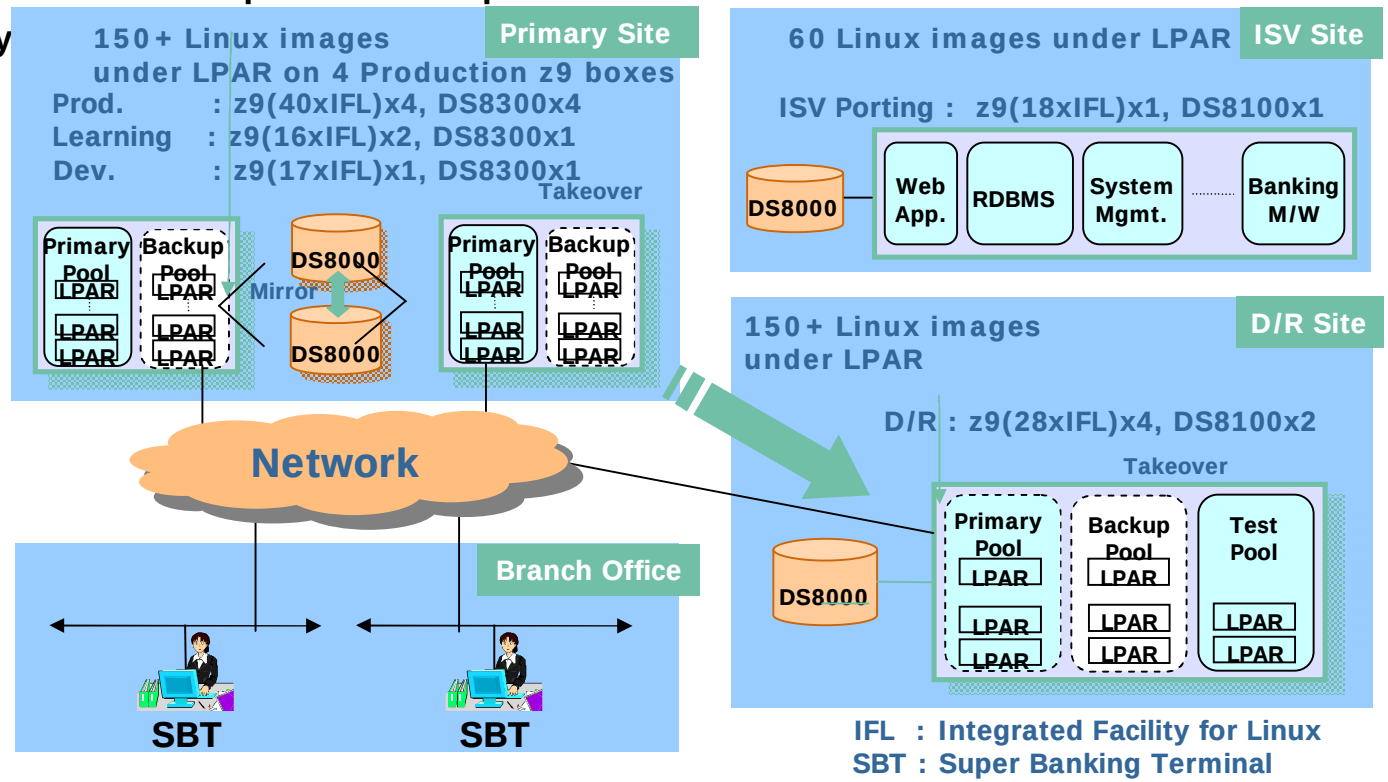
Customer Issues & Requirements:

Japanese Bank

Branch servers were based on Windows NT and reached the end of support

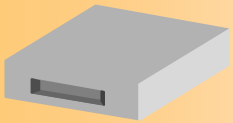
- About 800 servers constantly causing troubles
- Lower TCO by consolidating servers on a stable & reliable system
- Core banking applications needed to be ported to new platform
- Needed Disaster Recovery

Solution Overview



Japanese Bank

762 IA Server



386.3KW

Power consumption per hour

1878t

Annual CO2 Emission

6210 trees

Required trees to absorb CO2

Green Effect

4 System z9 EC

88.2kW

Power consumption per hour

429t

Annual CO2 Emission

1418 trees

Required trees to absorb CO2

The Role of Linux in IBM Products

MCP based (Embedded Linux)

- True Embedded Devices (Controllers/Service Modules)
  - OS burned into Flash/ROM at manufacturing
  - System control/service stack must be operational at first boot
  - *Examples: FSP (System i/p/z), AMM (BladeCenter), IMM(System x)*

- Systems Management Devices
  - OS needs to be pre-installed
  - Management stack must be operational immediately
  - *Examples:HMC (Power & z), System z Service Element, SanVC (Storage)*

- Special Purpose Appliances
  - OS and software stack combination manufactured into device
  - Customer cannot install OS after system purchase
  - *Examples: RSS 4690(RSS), DataPower(SWG), Image Capture(GBS), XIV((Storage)*

- Diagnostics/Systems Deployment
  - Diagnostic image delivered as bootable CD, flash drive
  - Image cannot be created by customer to include OS and diagnostics
  - *Examples: ToolsCenter(System x), RSS Diags(RSS), Tivoli OS Provisioning(SWG)*

Software Group Offerings

- OpenClient for Linux
- IBM Client for Smart Work
- Websphere Cloud Burst

Vision

on a Chip

to a Solution Center

in one System

From a Data Center