Virtualization Approaches for ISS Servers beyond Virtual Machine Technology

GI Fachgruppe Betriebssysteme
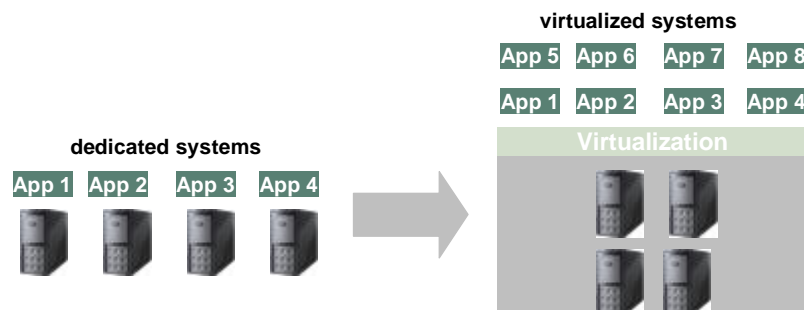
Guenther Aust   October 2006

We make sure | FUJITSU COMPUTERS SIEMENS

How we define Server Virtualization

n Virtualization separates OS, applications and data from dedicated systems

n pooled and shared resources for a flexible, dynamic and efficient usage

n IT can adapt quickly to changing requirements gaining business agility

n better utilization of invested resources improves business efficiency of IT

virtualized systems

App 5   App 6   App 7   App 8

App 1   App 2   App 3   App 4

Virtualization

dedicated systems

App 1   App 2   App 3   App 4

The initial Need for Virtualization

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

The Problem:
Hardware resource utilization

	Peak-hour Utilization	Prime-shift Utilization	24-hour Period Utilization
Mainframe	85-100%	70%	60%
Unix	50-70%	10-15%	<10%
Intel-based	30%	5-10%	2-5%
Storage	N/A	N/A	52%

Source: IBM Scorpion Whitepaper, Simplifying Corporate IT Infrastructure, 2000

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

What makes the difference?
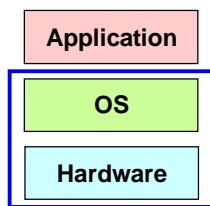
Traditional | Virtual machine
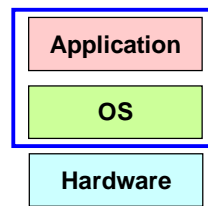
Application

OS

Hardware

Application

OS

Hardware

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved
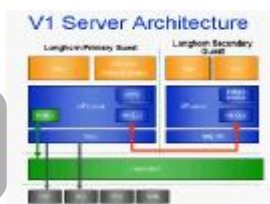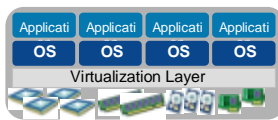

Different VMM Approaches
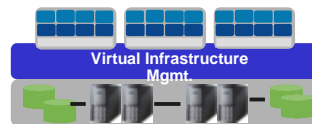
MS Hypervisor 2008

VMware ESX

Open Source XEN

V1 Server Architecture

| Applicati | Applicati | Applicati | Applicati |
| OS | OS | OS | OS |
| Virtualization Layer |

Virtual Infrastructure Mgmt.

Virtual Infrastructure Management

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

What Customers expect from Virtual Machine Concepts

Large enterprises tend to focus on the cost savings associated with virtualization

1. **Consolidation** — This represents dramatic capital equipment savings for most large enterprises.

2. **Deployment** — This reduces the administrative burden of loading and configuring the seemingly endless number of new servers arriving on the loading dock.

3. **Agility** — Virtualization reduces the administrative burden of moving workloads from one server to another to address new and changing demands.

4. **Freedom of choice** — This allows large enterprises to more easily accommodate a mix of vendors and server models within the data center.

5. **Protection** — This simplifies disaster recovery, but only for x86 platforms. Due to heterogeneity, most large enterprises have a more complex infrastructure to replicate in the case of a disaster

What Customers expect from Virtual Machine Concepts

Midsize businesses, while also driven by the promise of consolidation and cost savings, see virtualization chance to enable solutions what would otherwise be difficult and expensive

1. **Consolidation** —The level of savings is lower because of the scale of server deployments. By 2007, 40 percent of midsize businesses will reduce their server population by at least 20 percent through

2. **Protection** — virtualization is coupled with low-cost SAN solutions. The cost and complexity of implementing disaster recovery is reduced.

3. **Deployment** — Most midsize businesses have limited administrative resources; virtualization for them provides less effort and greater speed.

4. **Freedom of choice** — Virtualization allows midsize enterprises purchases based on competitive pricing without worrying about the overhead of supporting multiple vendors

5. **Agility** — Virtualization helps midsize businesses adapt server resources to address changes in workload demands. not at the same level of large enterprises. It makes it easier to bring up new services in remote branch offices
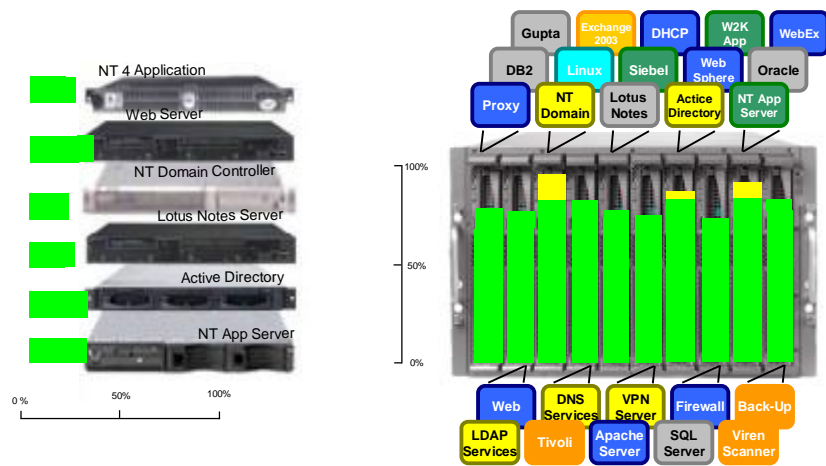
FUJITSU COMPUTERS
SIEMENS

Typical Usage Scenarios of Virtual Machines

Guenther Aust   Oktober 2006

Consolidate & virtualize for improved utilization and increased lifecycles off applications

We make sure

FUJITSU COMPUTERS
SIEMENS

NT 4 Application

Web Server

NT Domain Controller

Lotus Notes Server

Active Directory

NT App Server

0 %   50%   100%

Gupta Exchange 2003 DHCP W2K App WebEx
DB2 Linux Siebel Web Sphere Oracle
Proxy NT Domain Lotus Notes Actice Directory NT App Server

100%
50%
0%

Web DNS Services VPN Server Firewall Back-Up
LDAP Services Tivoli Apache Server SQL Server Viren Scanner
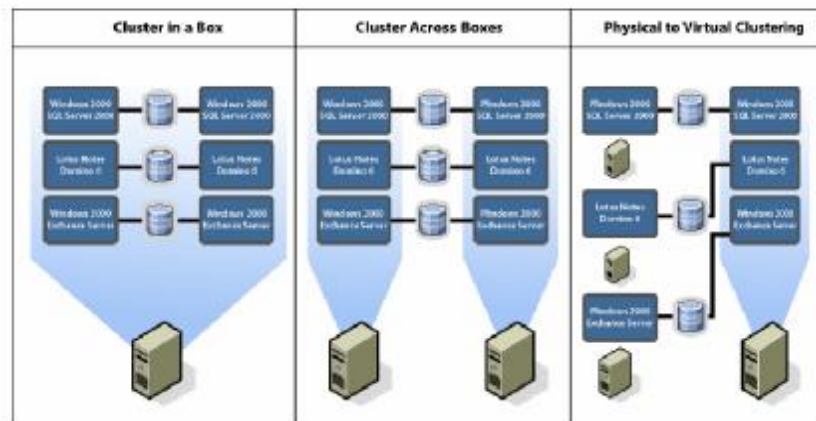
5

Deployment and Agility

- n Save Development Systems:
  Application Test
  on staging systems

- n Easy Deployment
  Copy and move
  Virtual Machines

- n Provide new machines
  within minutes

- n System Back-up copies
  for fast recovery

Production Systems

New York

Hong Kong

Tokyo

London

Development, QA, Test and Staging

Virtual Servers

Protection

Cluster in a Box | Cluster Across Boxes | Physical to Virtual Clustering

6

Protection and Agility

n High availability for all your VMs

n Impacted VMs are restarted on remaining hosts
 o Placement optimized by global scheduler

n None of the cost and complexity of clustering

n Creating a Unified Compute Resource. Global scheduler:
 o Automates initial virtual machine placement
 o Uses VMotion to optimize based on current workload
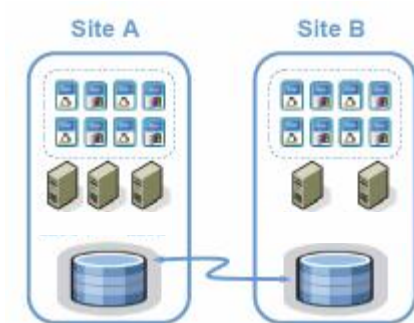 o Reacts to adding or removing hosts from the ESX server farm
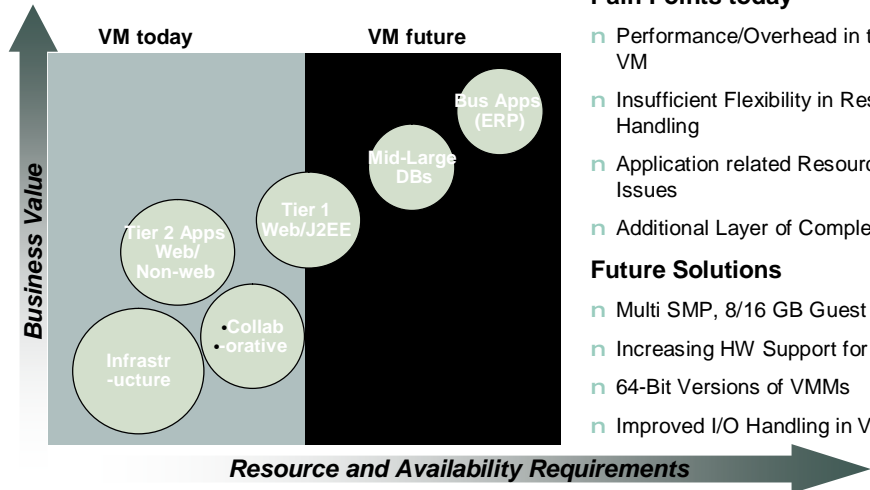
Protection and Freedom of Choice

Site A   Site B

n Implement DR scenarios across heterogeneous server environments

n Minimize hardware investment by running Active/Active configurations

n Failover services in both directions

n Grant high resource priority to critical applications in failover scenarios

n Future: DR between real and virtualized Infrastructures

What we have and what we miss

VM today

VM future

Business Value

- Bus Apps (ERP)
- Mid-Large DBs
- Tier 1 Web/J2EE
- Tier 2 Apps Web/Non-web
- Collaborative
- Infrastr-ucture

Resource and Availability Requirements

Pain Points today

- Performance/Overhead in the indiv. VM
- Insufficient Flexibility in Resource Handling
- Application related Resource Issues
- Additional Layer of Complexity

Future Solutions

- Multi SMP, 8/16 GB Guest VM
- Increasing HW Support for VMMs
- 64-Bit Versions of VMMs
- Improved I/O Handling in VMMs

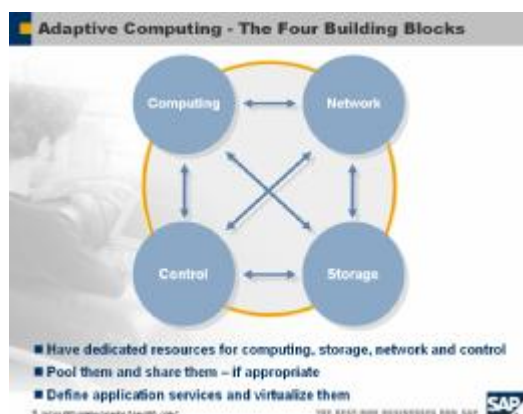# Application Virtualization

FlexFrame for mySAP und ORACLE 10g

FlexFrame – based on a joint research project with SAP and NetApp

Fujitsu Siemens Computers and SAP work actively together as partners in the area of Adaptive Computing.
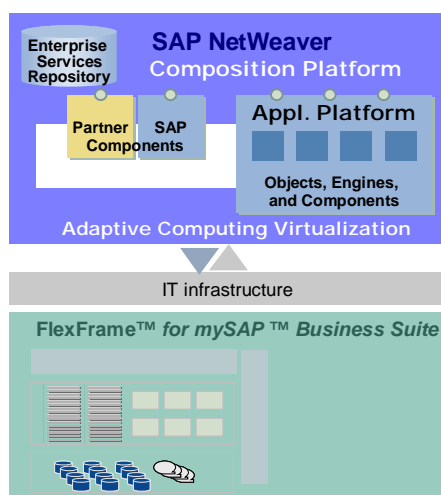
With FlexFrame™ for mySAP Business Suite™ Fujitsu Siemens Computers is an early adopter of the Adaptive Computing capability.

As a consequence, they are the first to sell this infrastructure solution and have first productive installations.

Quote: SAP Service Marketplace – FAQ, Adaptive Computing

FlexFrame™ is the IT-Infrastructure for SAP NetWeaver™

SAP NetWeaver with Adaptive Computing

§ **Virtualization of application services**
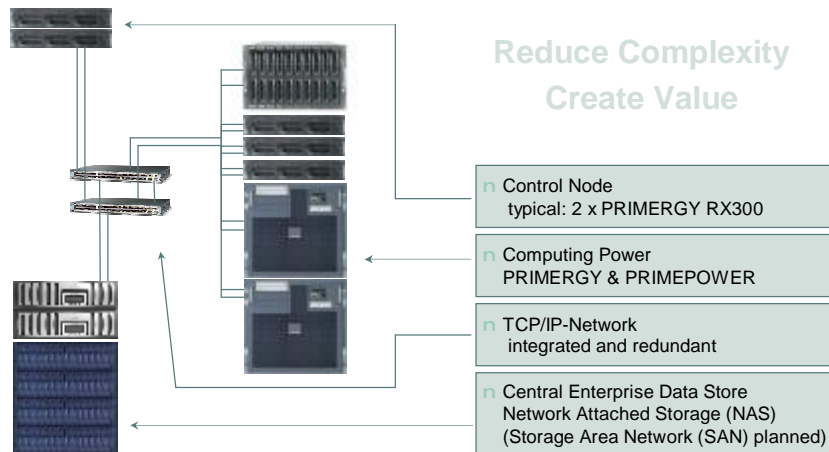
FlexFrame – the DDC solution based on

§ **Virtualization**
§ **Automation**
§ **Integration**

to add further benefits

§ **Cost efficiency**
§ **High Availability**
§ **Easy & Flexible**

FlexFrame – Building Blocks

Reduce Complexity
Create Value

n Control Node
   typical: 2 x PRIMERGY RX300

n Computing Power
   PRIMERGY & PRIMEPOWER

n TCP/IP-Network
   integrated and redundant

n Central Enterprise Data Store
   Network Attached Storage (NAS)
   (Storage Area Network (SAN) planned)

FlexFrame Virtualization

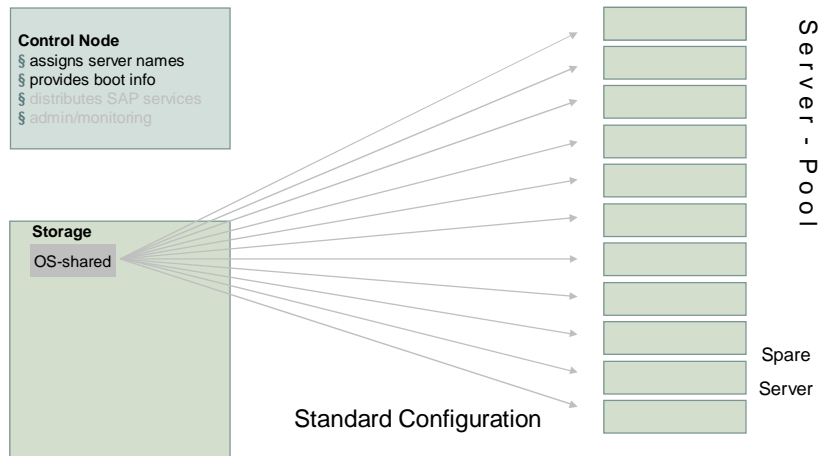Virtualization on several levels is one of the most important features within FlexFrame to increase flexibility and to drive down costs

n Shared Operating System
   → less admin efforts and no provisioning issues

n Virtualization of SAP Services
   → leads to higher utilization and easy server changes

n Virtual LAN technology
   → less complexity in cabling, simple & easy to manage, sophisticated business continuity

n Virtual storage concept
   → storage grid
   → highest flexibility for your data while providing the best performance

10

The FlexFrame concept – the shared operating system

FUJITSU COMPUTERS
SIEMENS

Control Node
§ assigns server names
§ provides boot info
§ distributes SAP services
§ admin/monitoring

Storage
OS-shared

Server - Pool

Spare

Server

Standard Configuration

FlexFrame Concept – Virtualized SAP Services

FUJITSU COMPUTERS
SIEMENS

Control Node / ACC
assigns server names
provides boot info
distributes SAP services
admin/monitoring

OS-shared

Enterprise Portal

ECC

Business Warehouse

SAP J2EE

SRM

Standard-Configuration

FlexFrame Concept – Virtualized SAP Services

FUJITSU COMPUTERS
SIEMENS

Control Node / ACC
assigns server names
provides boot info
distributes SAP services
admin/monitoring

OS-shared

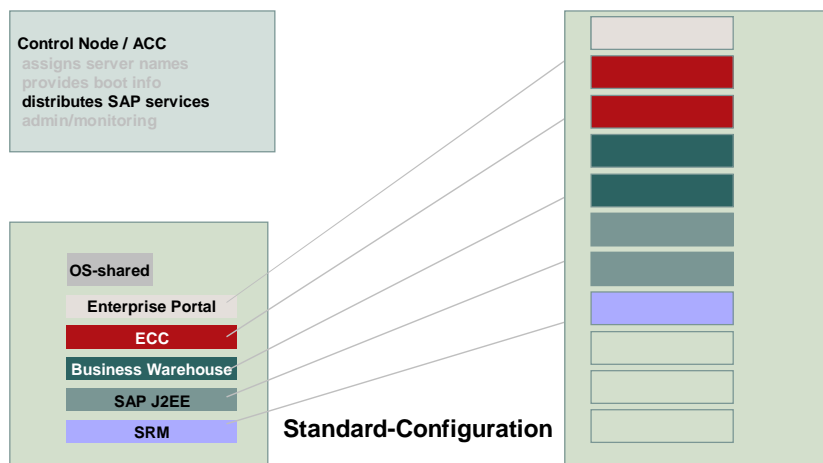| Enterprise Portal |
| ECC |
| Business Warehouse |
| SAP J2EE |
| SRM |

additional

for peak handling

Login in the morning

23   Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

---

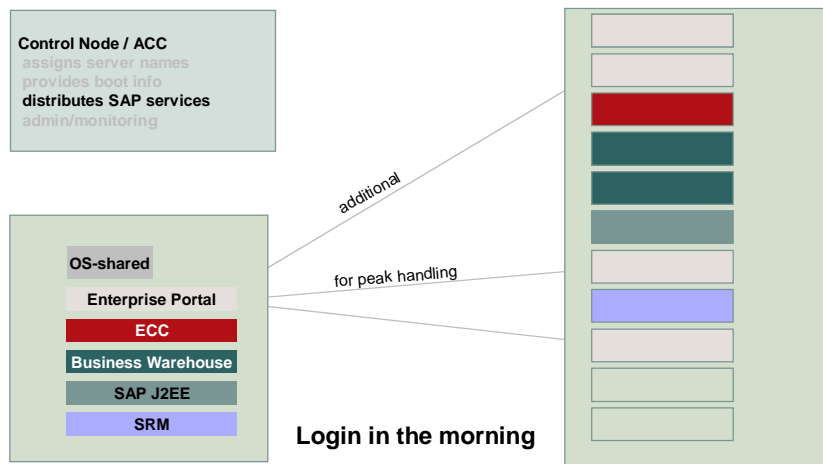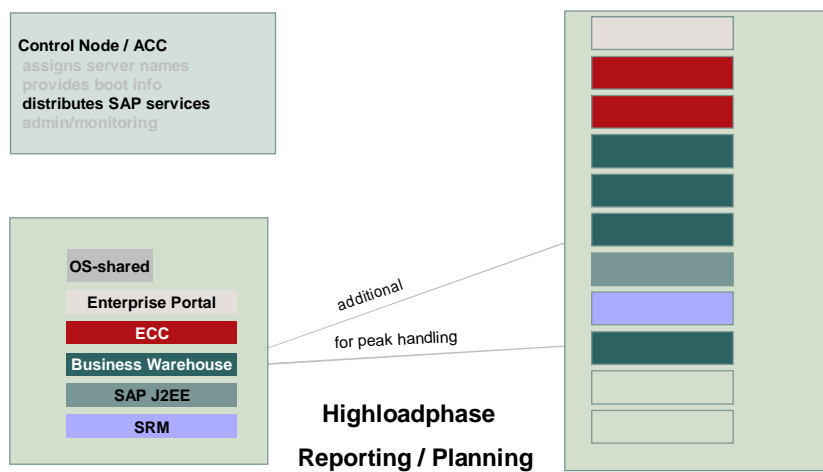## FlexFrame Concept – Virtualized SAP Services

FUJITSU COMPUTERS
SIEMENS

Control Node / ACC
assigns server names
provides boot info
distributes SAP services
admin/monitoring

OS-shared

| Enterprise Portal |
| ECC |
| Business Warehouse |
| SAP J2EE |
| SRM |

additional

for peak handling

Highloadphase

Reporting / Planning

24   Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved
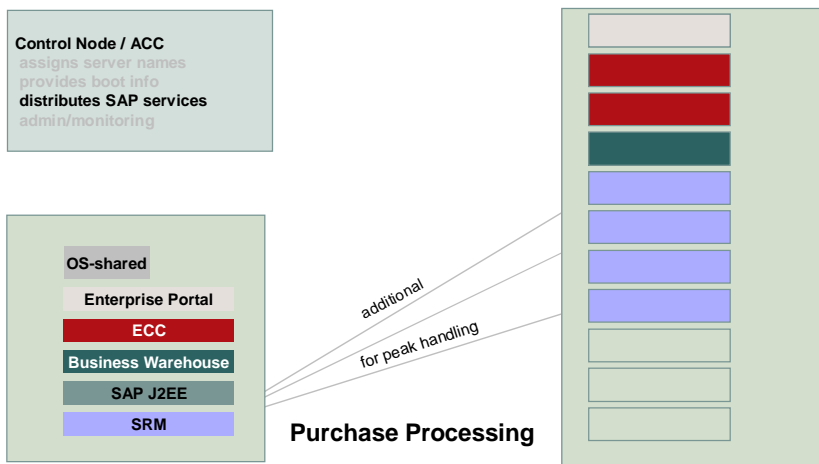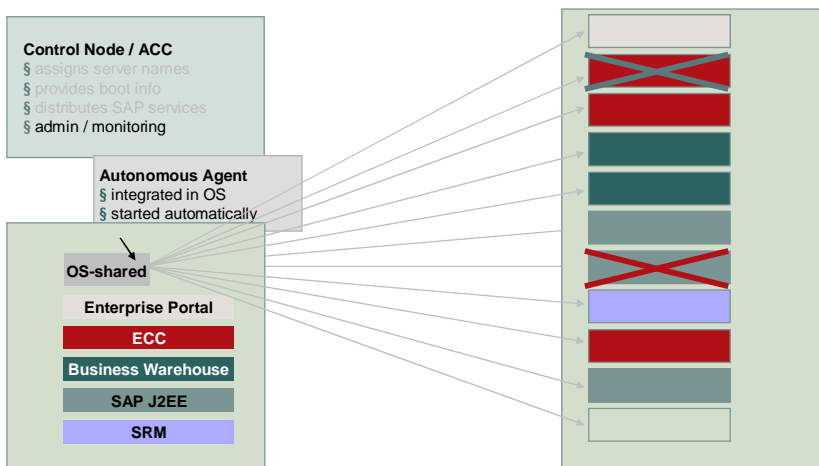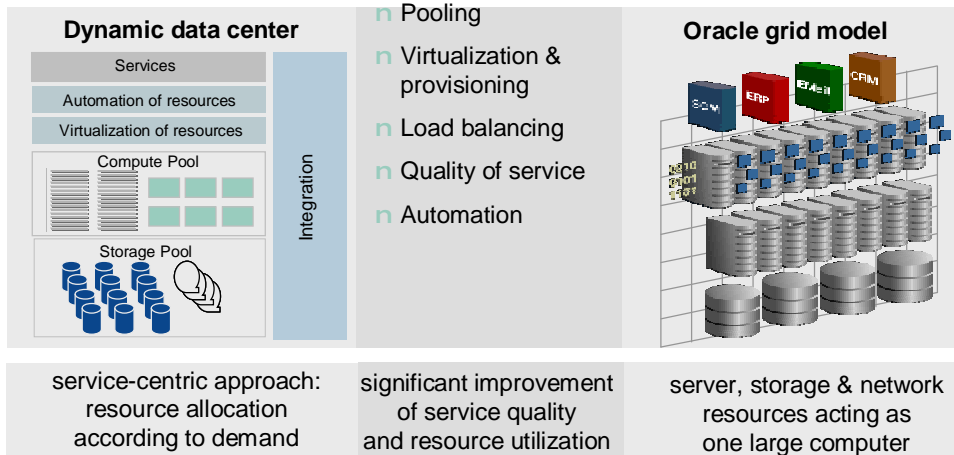
12

FlexFrame Concept – Virtualized SAP Services

Control Node / ACC
assigns server names
provides boot info
distributes SAP services
admin/monitoring

OS-shared

Enterprise Portal

ECC

Business Warehouse

SAP J2EE

SRM

additional

for peak handling

Purchase Processing

FlexFrame – High Availability for all SAP Services

Control Node / ACC
§ assigns server names
§ provides boot info
§ distributes SAP services
§ admin / monitoring

Autonomous Agent
§ integrated in OS
§ started automatically

OS-shared

Enterprise Portal

ECC

Business Warehouse

SAP J2EE

SRM

13

FlexFrame for Oracle – joint project with partner Oracle

Joint vision

Dynamic data center

- Services
- Automation of resources
- Virtualization of resources
- Compute Pool
- Storage Pool

Integration

n Pooling

n Virtualization & provisioning

n Load balancing

n Quality of service

n Automation

Oracle grid model

| service-centric approach: resource allocation according to demand | significant improvement of service quality and resource utilization | server, storage & network resources acting as one large computer |

Solution elements – software

n Operating system

  n Linux (SUSE and RedHat)

n Oracle core technologies

  n Database, real application cluster

  n Application Server (optional)

  n Monitoring (Grid Control)

  ORACLE DATABASE 10*g* ORACLE APPLICATION SERVER 10*g*
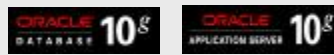
n Automation technologies

  n Adaptive Services Control Center

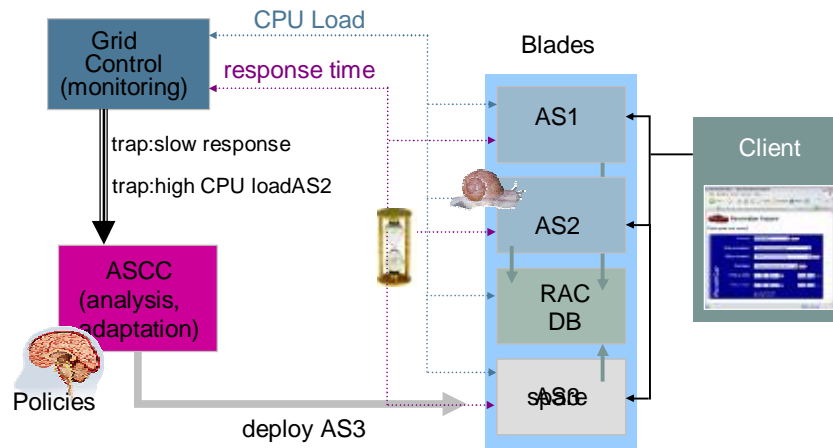 **n** Drives autonomic cycles (e.g. workload mgmt and failover) based on defined policies

  AS CC
  Analysis — Adaptation
  knowledge & policy
  Monitoring — Execution

n Virtualization and provisioning technologies

  n Cloning of operating system, database and application server

  n Dynamic provisioning via netboot

Sample scenario: policy-based workload management

CPU Load

Grid
Control
(monitoring)

response time

Blades

AS1

Client

trap:slow response

trap:high CPU loadAS2

AS2

ASCC
(analysis,
adaptation)

RAC
DB

Policies

AS3
space

deploy AS3

PRIMERGY BladeFrame

Virtualization thru stateless Servers
I/O Virtualization in switched Fabric
Reduced Complexity

Guenther Aust   October 2006

ISS Servers – Quo vadis?
Typical Blade Platform Today

- Current models are typically 6U to 7U chassis with 10 to 14 1P/2P x64 blades
- Each blade is like a server motherboard
  - IDE/SCSI attached disks, network and IO Daughter card on the blade
  - Midplane is passive; routing is very complex; IO switches provided in the chassis
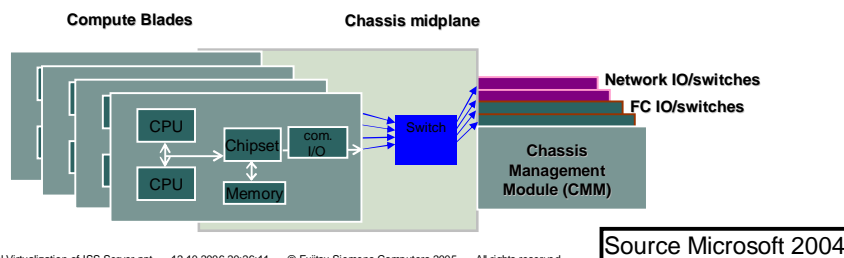 - SAN attached rate is high, ~40%
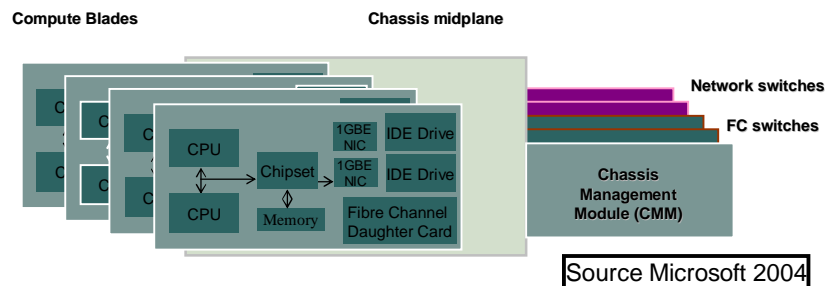
Source Microsoft 2004

ISS Servers – Quo vadis?
Future Blade Platform

- Similar chassis configuration, e.g. 6U to 7U chassis with 10 to 14 1P/2P x64 blades
- The compute blade becomes stateless
  - All IO and direct attached disks are removed
  - Consolidated storage on FC or iSCSI SAN
  - Simplifies blade failover and repurposing
- The chassis contains a set of configurable components
  - All IO devices and switches are at the far end of the midplane
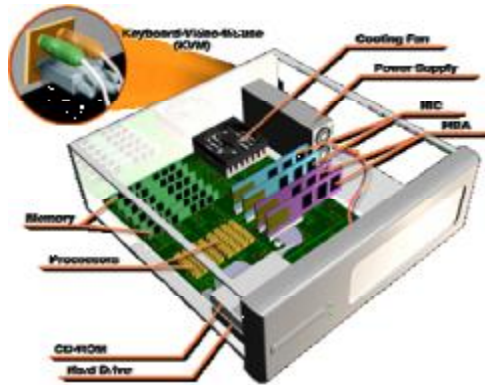- The server becomes stateless

Source Microsoft 2004

Commodity technology alleviates the problem of expensive CPU and Memory

- **n** **Intel / AMD price / performance**
- **n** **Platform independent operating systems**
- **n** **Lower capital cost for compute resources**

BUT what surrounds the CPU and memory is complex and expensive

- **n** Particularly challenging in high-end

- **n** Infrastructure team must piece together and maintain
  - o Multiple I/O connections
  - o Multiple switches
  - o Complex management software
  - o Clustering software
  - o Multi-path software
  - o Provisioning software
  - o Load balancers
  - o Multiple software licenses
  - o Complex DR and DR verification processes
  - o etc….

Legacy Blade Architecture as of today

A Fundamental Re-design by replacing NICs/ HBA's, local CDROM's, local disks, local KVM
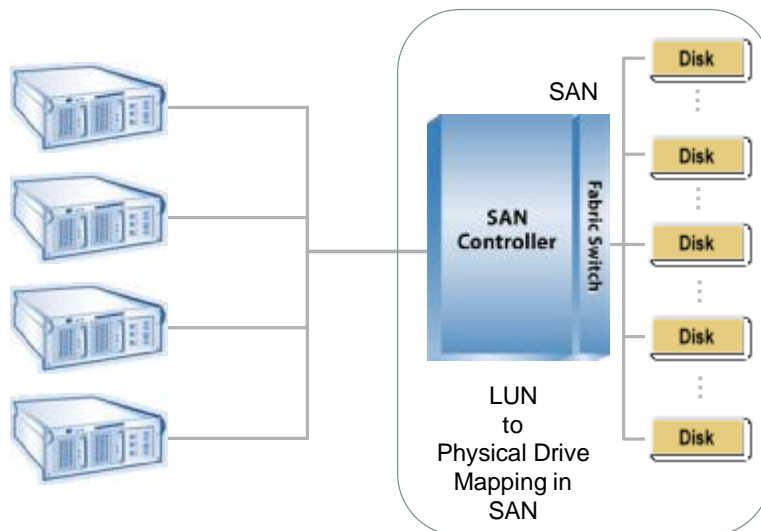

pBlade

consistent form factor
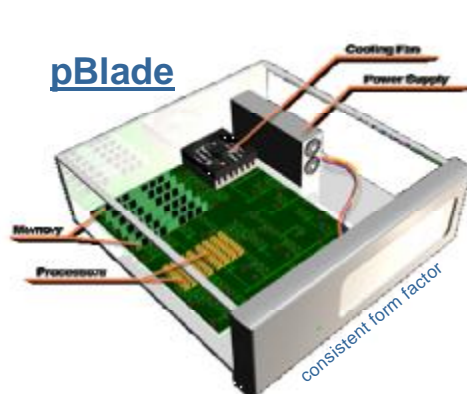
Simplicity leads to:

- n less failure rate
- n no overheating
- n less power consumption
- n real processing power flexiblity

18

Inspiration was the SAN

SAN

SAN Controller

Fabric Switch

Disk
Disk
Disk
Disk
Disk

LUN
to
Physical Drive
Mapping in
SAN

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

Processing Area Network (PAN) Architecture

CPU Memory
CPU Memory
CPU Memory
CPU Memory
CPU Memory

PAN

Fabric Switch

PAN Controller

SAN

SAN Controller

Fabric Switch

Disk
Disk
Disk
Disk
Disk

Logical Server
Definition Mapping
to
Physical CPU's,
Memory & I/O

LUN to
Physical drive
Mapping in SAN

Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved
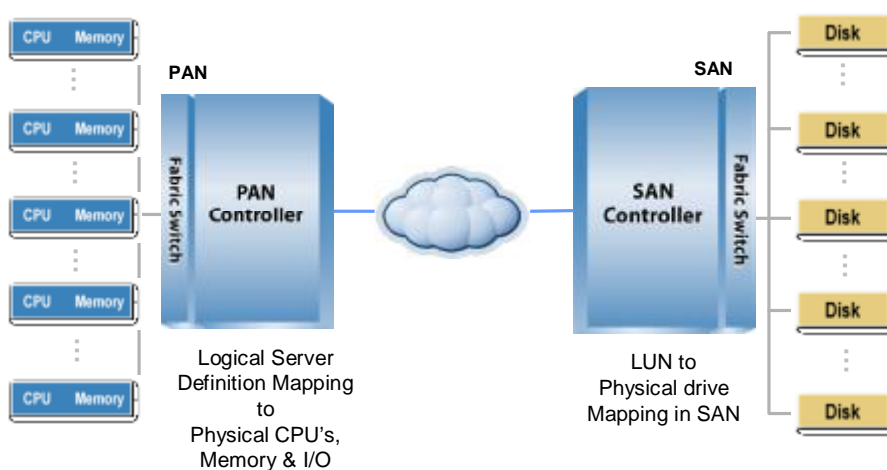
High-speed fabric connects anonymous pBlades
## PAN controller hosts PAN OS and bridges physical I/O

Fabric Switch

- **n** High speed/low latency to support combined disk and network traffic
- **n** High performance for mission critical, network intensive apps

PAN Controller

- **n** Stores server definitions
- **n** Automated management
- **n** Logical resource configuration
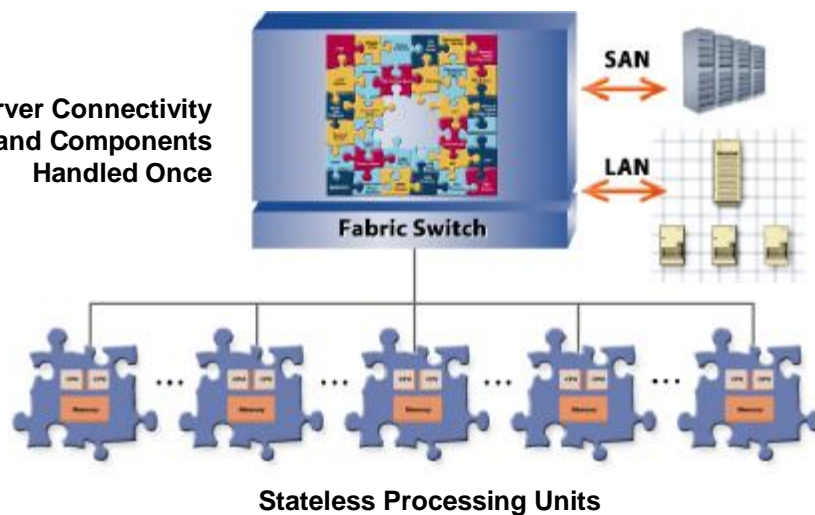- **n** Contains standard NICs and HBAs
- **n** Consolidate I/O

39   Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved

---
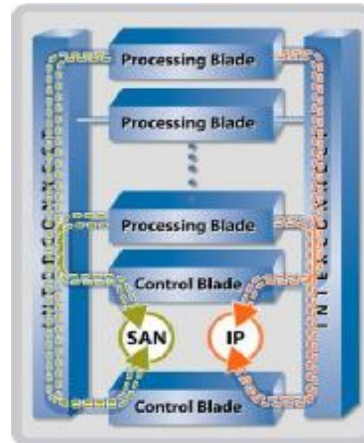
Processing Area Network Architecture

Server Connectivity and Components Handled Once

Stateless Processing Units

40   Final Virtualization of ISS Server.ppt   12.10.2006 20:26:11   © Fujitsu Siemens Computers 2005   All rights reserved
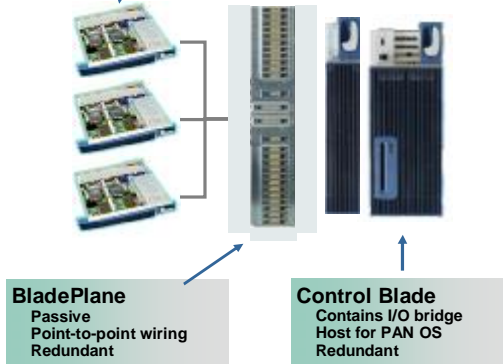
Putting it all together

Processing Blades
Single form factor
High performance
Stateless and anonymous

Switch Blade
High-speed, point- to-point network
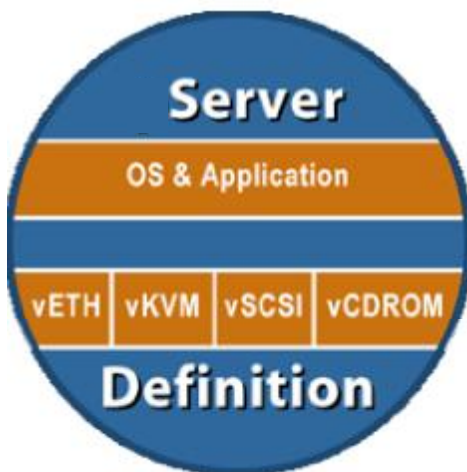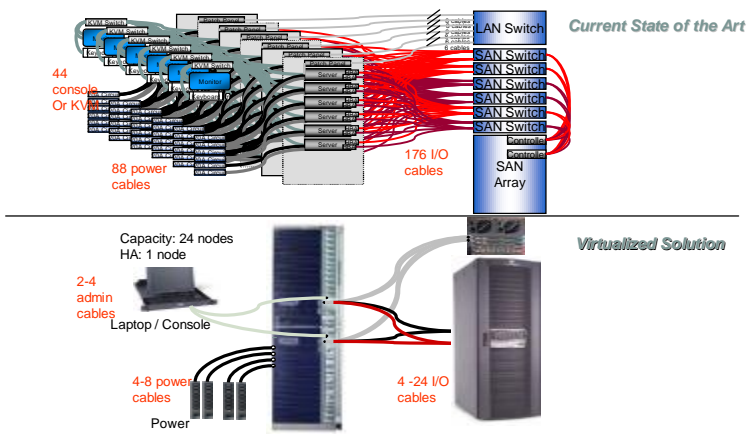Invisible outside frame
Redundant

BladePlane
Passive
Point-to-point wiring
Redundant

Control Blade
Contains I/O bridge
Host for PAN OS
Redundant

Provisioning More Than Just Software Using Virtualised Resources

n   End to End provisioning in minutes

Reduce Complexity

… plus dramatically reduced complexity

Current State of the Art

LAN Switch

SAN Switch
SAN Switch
SAN Switch
SAN Switch
SAN Switch
SAN Switch

Controller
Controller
SAN
Array

44 console Or KVM

88 power cables

176 I/O cables

Virtualized Solution

Capacity: 24 nodes
HA: 1 node

2-4 admin cables

Laptop / Console

4-8 power cables

Power

4 -24 I/O cables

Inherent capabilities of PAN
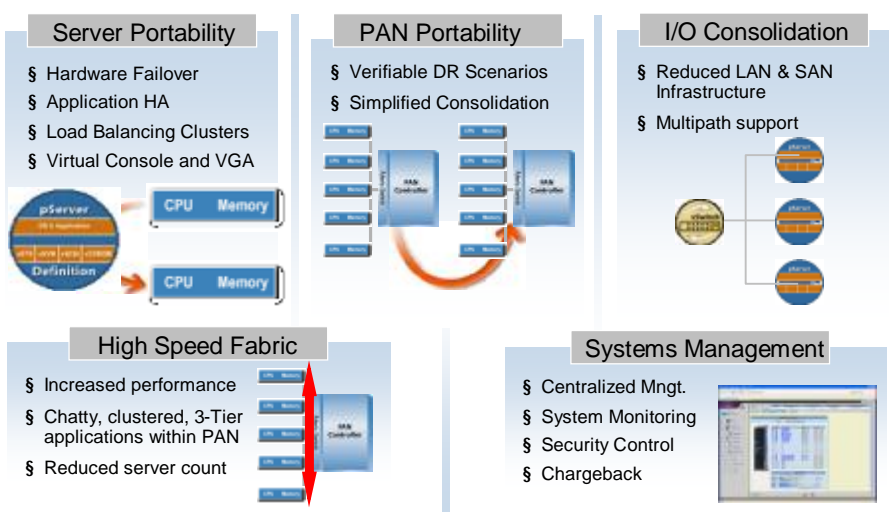
Server Portability

§ Hardware Failover
§ Application HA
§ Load Balancing Clusters
§ Virtual Console and VGA

PAN Portability

§ Verifiable DR Scenarios
§ Simplified Consolidation

I/O Consolidation

§ Reduced LAN & SAN Infrastructure
§ Multipath support

High Speed Fabric

§ Increased performance
§ Chatty, clustered, 3-Tier applications within PAN
§ Reduced server count

Systems Management

§ Centralized Mngt.
§ System Monitoring
§ Security Control
§ Chargeback

Different workloads over a full day (Illustration)
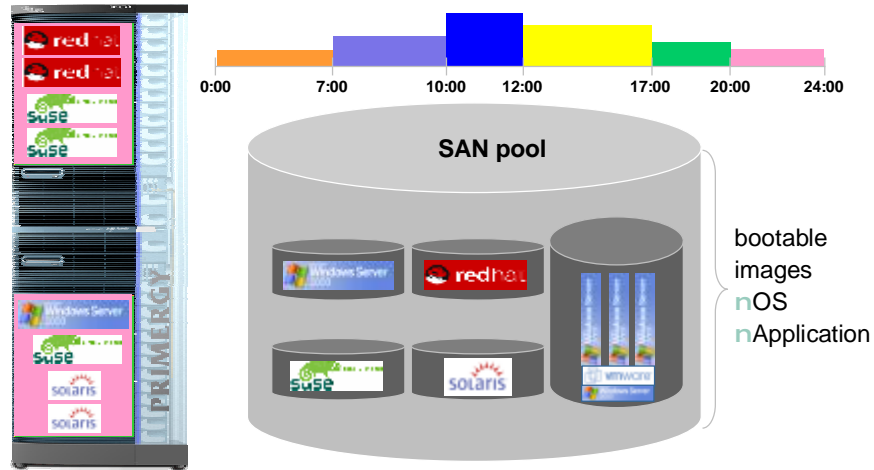
SAN pool

bootable
images
nOS
nApplication

0:00　7:00　10:00　12:00　17:00　20:00　24:00

IT with a sense of responsibility
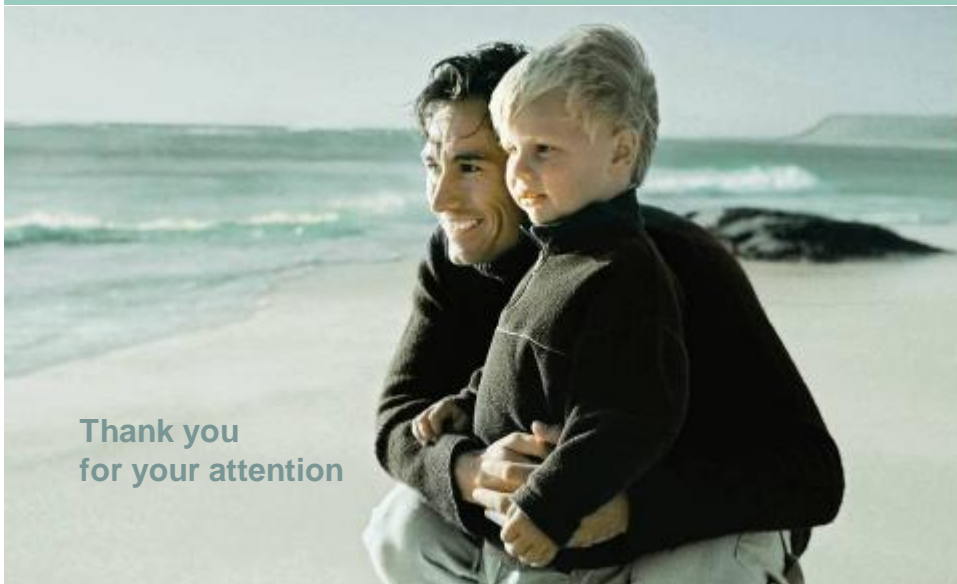
**Thank you
for your attention**