Cluster 2002

Blades—An Emerging System Design Model for Economic Delivery of High Performance Computing

Kirk Bresniker

Systems Architect
Internet and Applications Systems Laboratory

Hewlett-Packard

September 24, 2002

Cluster 2002

Agenda

Agenda

Introduction / Definitions

Origins of the Blade Server

Limitations of First Generation Designs

Second Generation Design

Future Directions

Cluster 2002

Definitions

A Blade is …

An inclusive computing system that includes processor, memory,network connections and associated electronics on a single motherboard.

The server blade typically is associated with an enclosure system that allows multiple blades to be housed in a standard server 'sub-rack' or enclosure that share resources such as power supplies and cooling fans.

Blades …

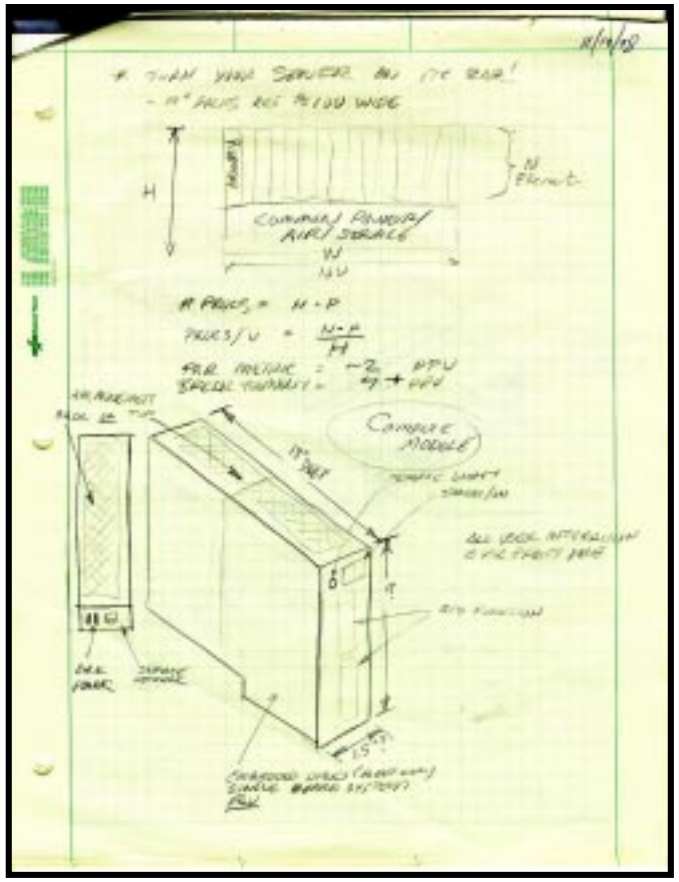
- are easily accessible
- offer increased computing density
- have modular architecture that ensures flexibility and scalability.

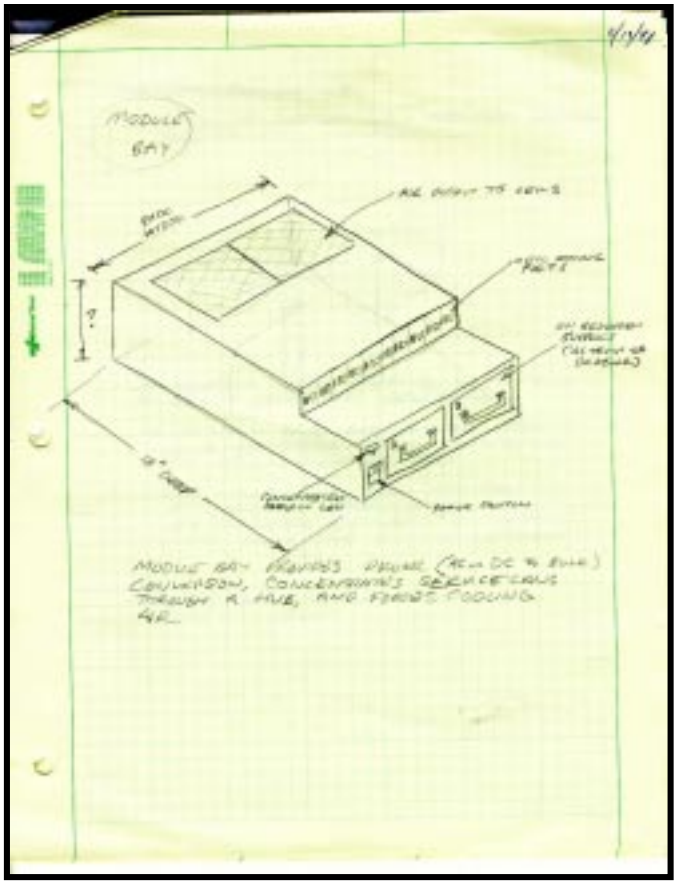
Source: Provisioning the Internet Infrastructure: Server Blades and Dynamic Workload Management IDC Document # 24155

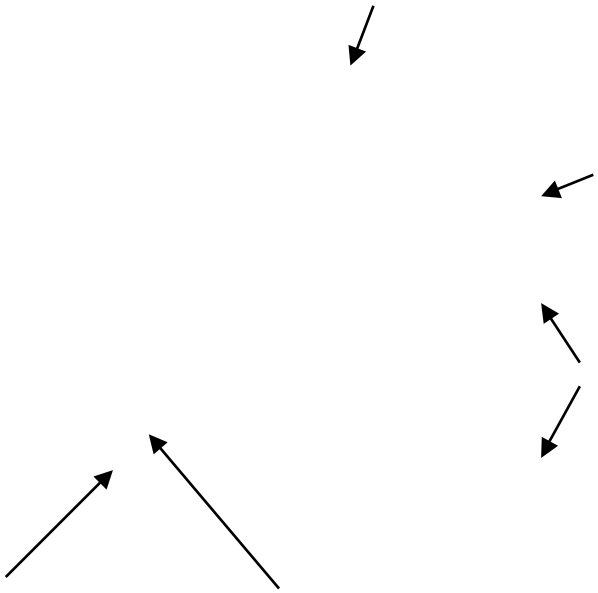
Cluster 2002

Origins

Initial Blade Sketches







Cluster 2002

Dominant xSP Concerns – Circa 1999

Origins

Time-to-Deployment - Competition is cut-throat. How quickly can a new service be deployed in the physical plant?

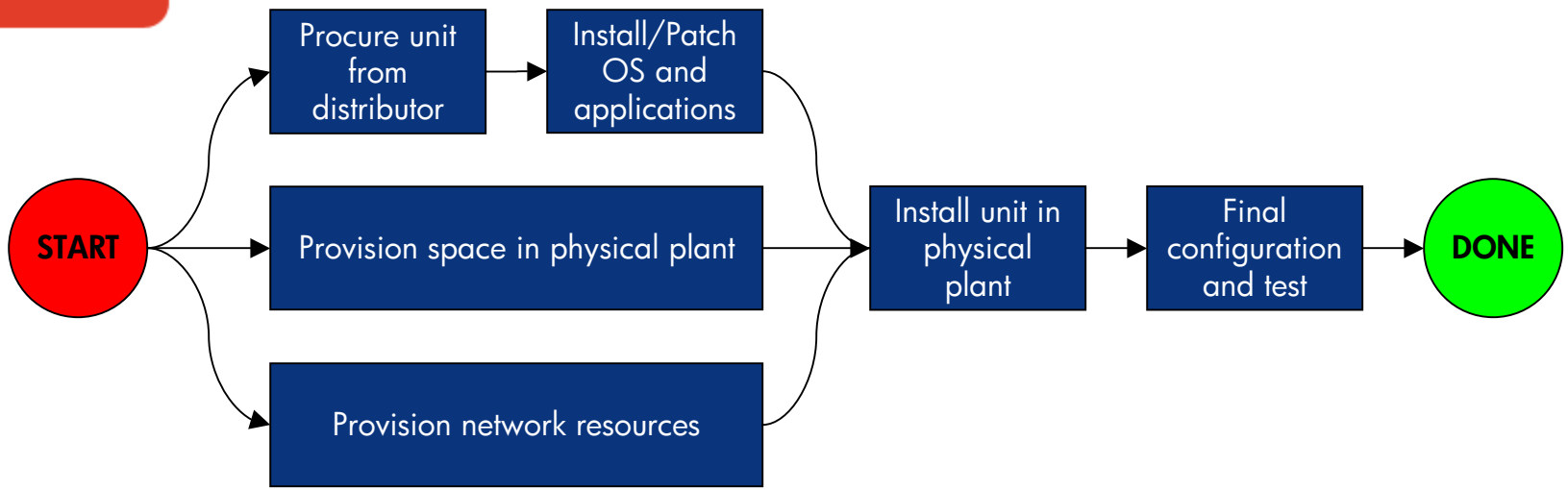
System Density - The data center incurs costs in $/area and recovers costs in systems/area. How many systems can be burdened on each square foot of data center?

Moving Customers up the price list - Currently shared hosting only sells at the introductory service level. Moving customers up the price list requires dedicated hardware

Cluster 2002

Origins

ASP Deployment Timeline – Circa 1999



Time to Deploy : 5 Days

Cluster 2002

Origins

Dominant xSP Concerns – Circa 1999

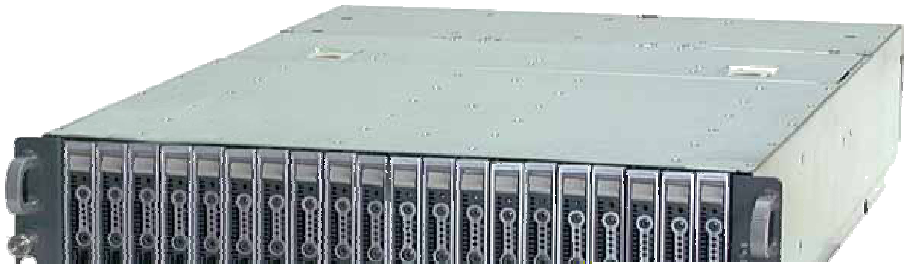
Time-to-Deployment - Competition is cut-throat. How quickly can a new service be deployed in the physical plant?

System Density - The data center incurs costs in $/area and recovers costs in systems/area. How many systems can be burdened on each square foot of data center?

Moving Customers up the price list - Currently shared hosting only sells at the introductory service level. Moving customers up the price list requires dedicated hardware

Cluster 2002

First Generation Blade



Each blade contains:

- 700MHz PIII 512K LV
- 512MB/1GB PC133 SDRAM
- 30GB Mobile IDE Drive
- Dual 10/100 NICs
- Integrated Management

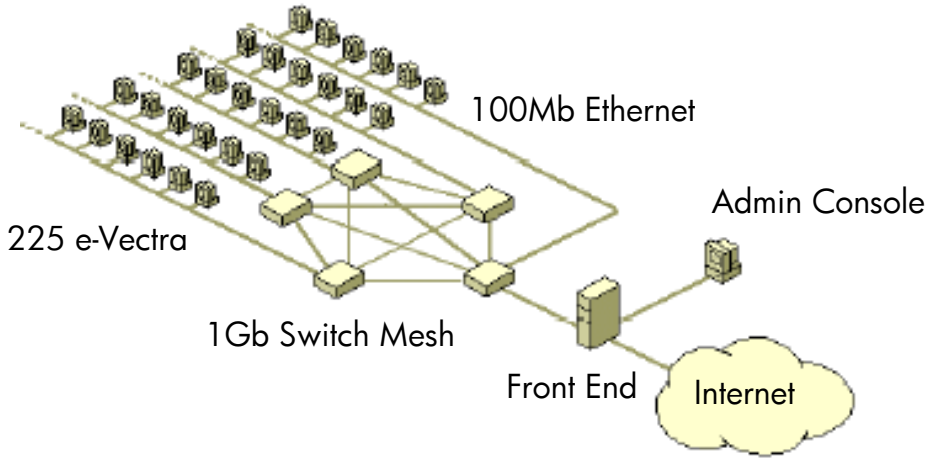
Each 3U chassis contains:

- 20 Server blades
- N+1 Power
- Dual Ethernet Switches

Cluster 2002

Mainstream Technology Clusters

First Generation



Cluster Topology

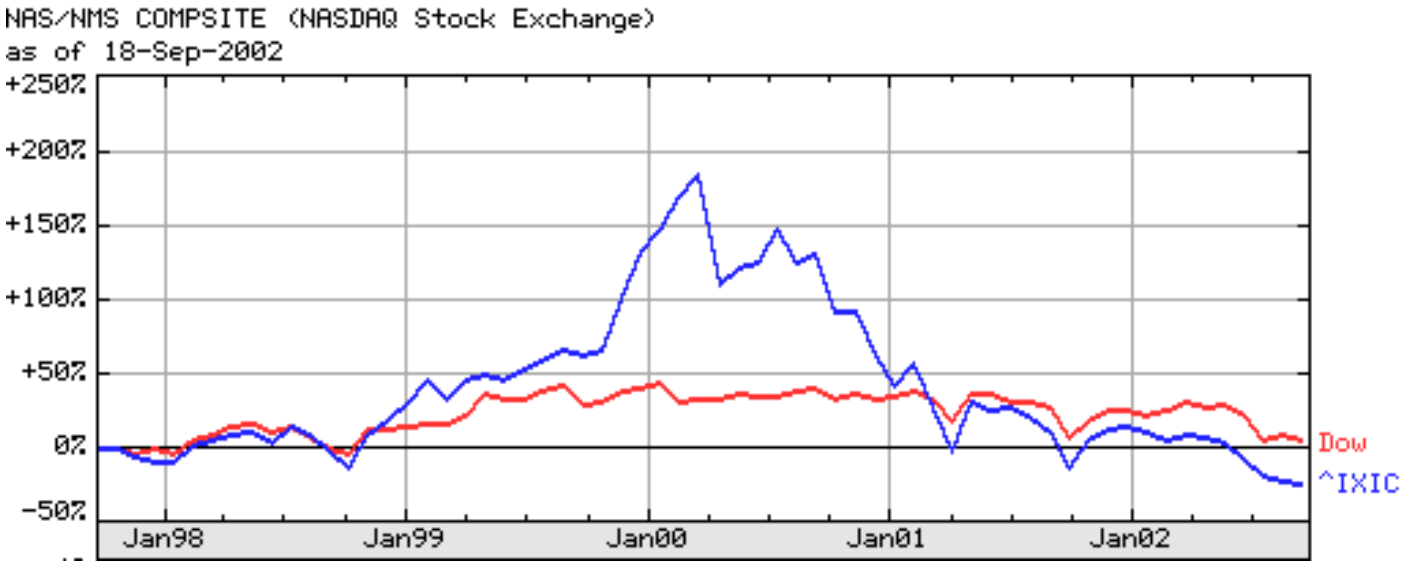
Cluster Racking

Mainstream Technology Cluster achieved rank of 385th on TOP500 as of 06/21/2001

Cluster 2002

Limitations

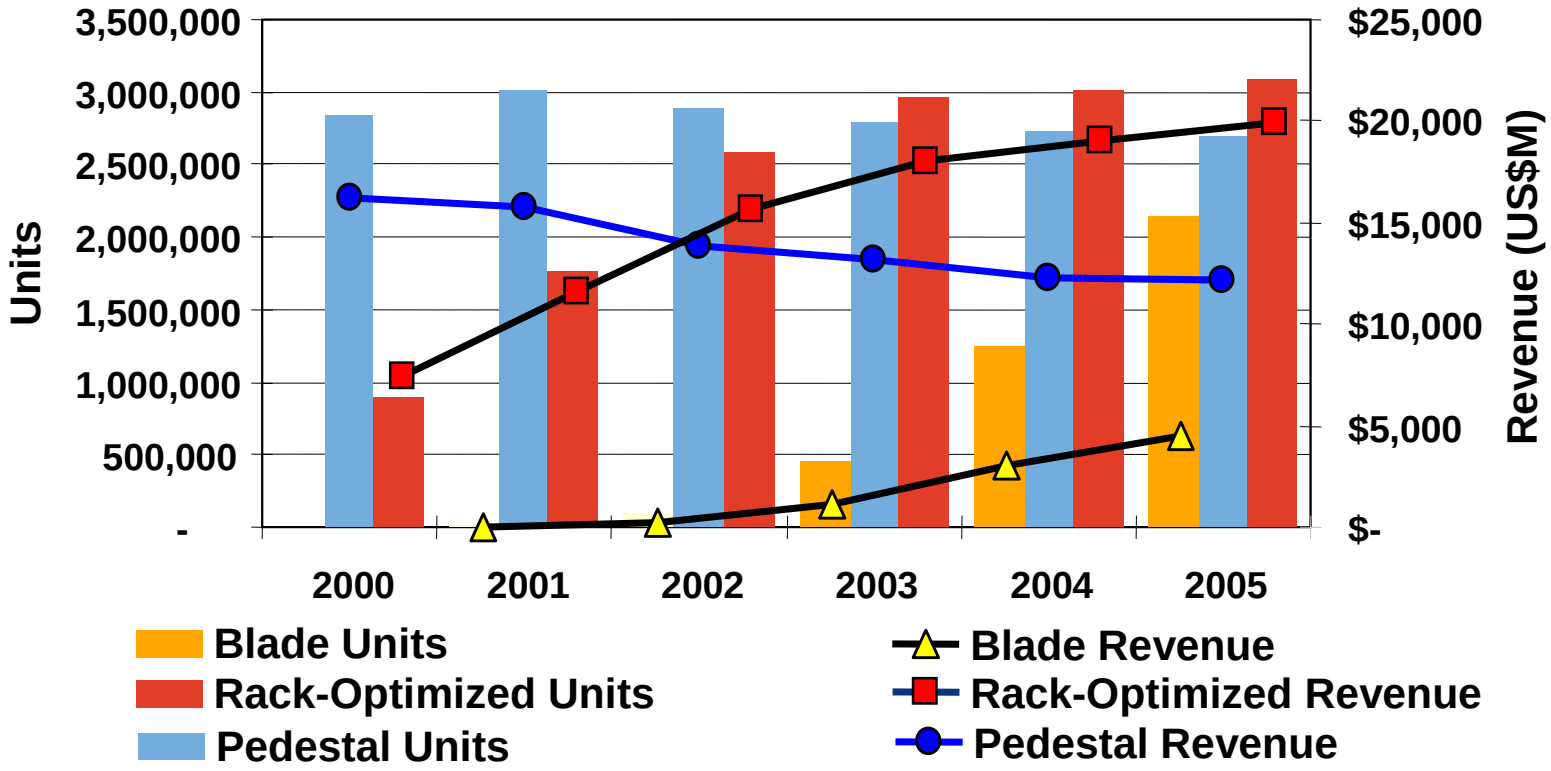
What New Economy?



Cluster 2002

Bladed Architectures: What's the Opportunity?

Second Generation



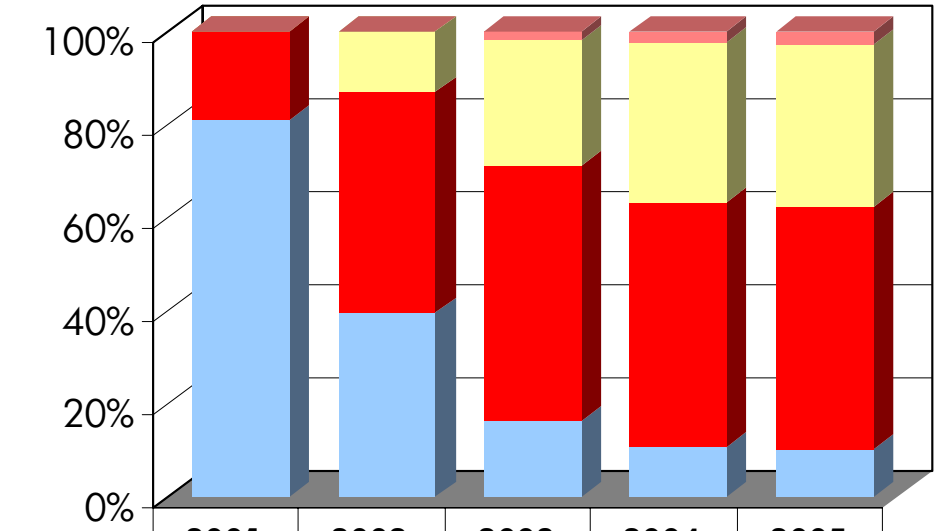
By 2005, IDC believes the blade form factor will capture approximately 27% of entry server unit sales and 12% of entry server revenue.

Source: IDC August 2001

Cluster 2002

Bladed Architectures: Where Are the Sweet Spots?

Second Generation



	2001	2002	2003	2004	2005
8-way	0%	0%	2%	2%	3%
4-way	0%	13%	27%	35%	35%
2-way	19%	47%	55%	52%	52%
1-way	81%	39%	16%	11%	10%

2005
$4.5 billion

Share of bladed server revenue quickly migrates to the 2 & 4 processor "flavors" as volume ramps and commoditization occurs

Source: IDC August 2001

Cluster 2002

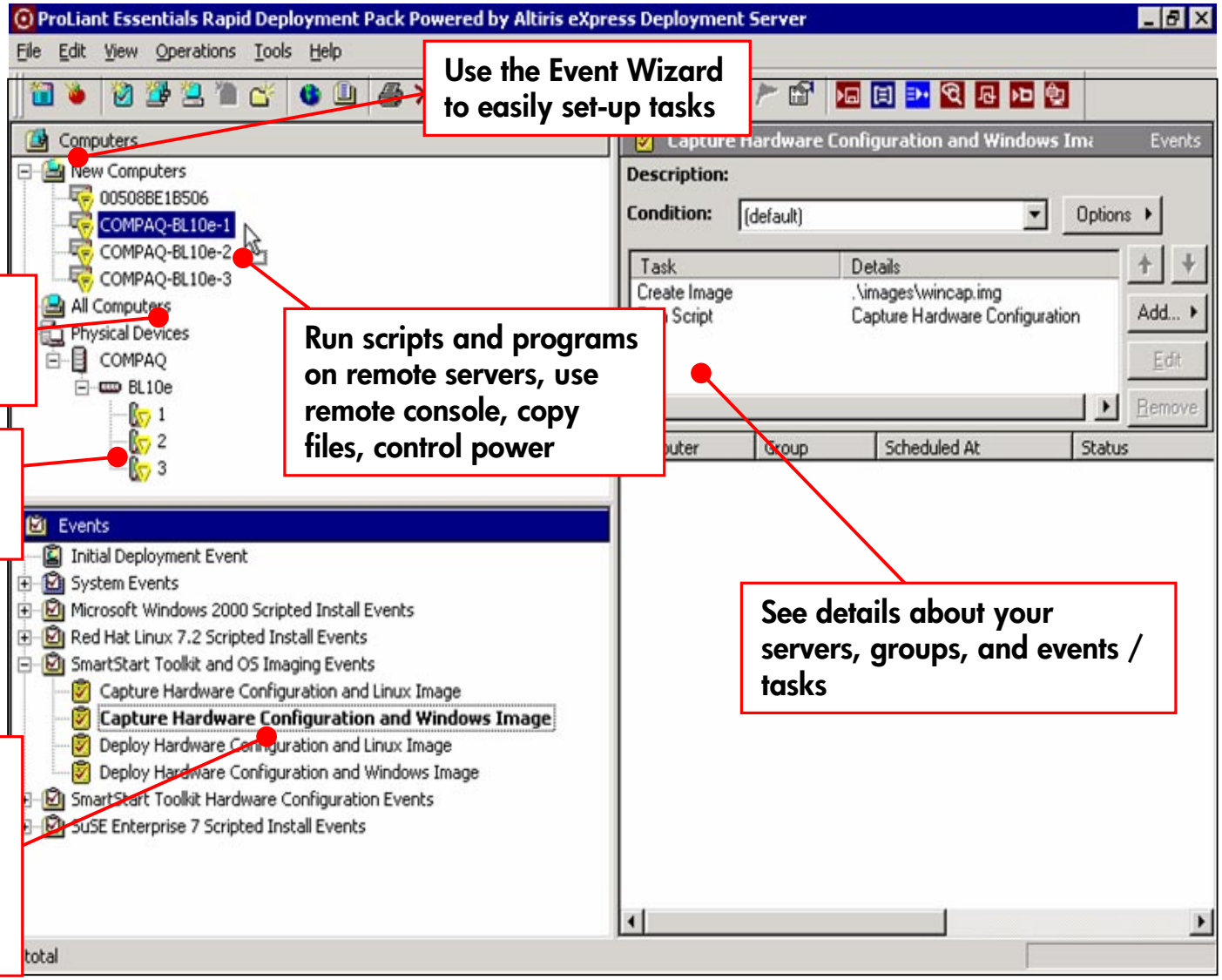
Second Generation Blade

- Processors: Mobile/LV/ULV → Mainline server Processors
- SMP: Single → Dual (or more) Processors
- Memory sizes: Sub-1GB per processors → Greater than 1GB per processor
- Ethernet NICs: Dual 10/100 → Multiple 10/100/1000
- Manageability: rack-optimized server leveraged → blade-aware
- Local Storage: low performance mobile → high performance, high reliability server storage
- Remote Storage: Low performance NAS → high performance NAS or SAN
- Chassis form factor: 3U → 6U
- Blades form factor increase in depth, width, and height

Cluster 2002

Second Generation

Rapid Deployment Tools



Flexible organization and grouping of all servers

View physical location of blade servers

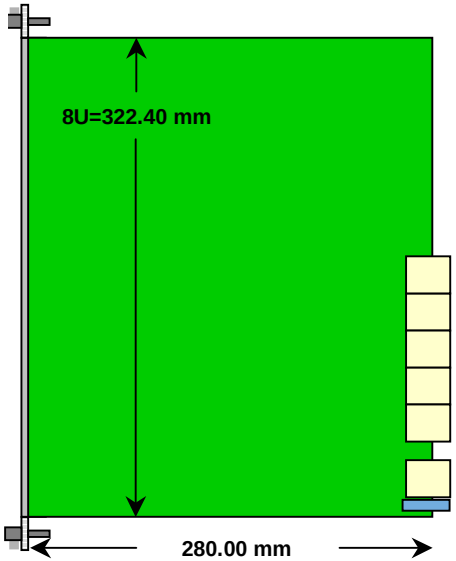
Run an event by simply dragging it to a server or group (or drag servers or groups to the event)

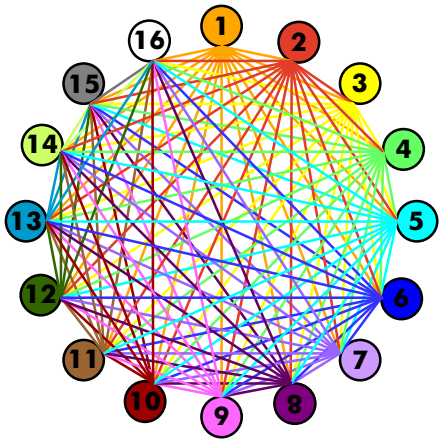
See details about your servers, groups, and events / tasks

Cluster 2002

Futures

Advanced Telecom Computing Architecture





Each blade contains:

- A 200W Thermal Envelope (enables Dual slot / DP IPF or QP IA-32)
- Fully connected Mesh Data Transport Protocol Agnostic Fabric. Each Link can support, for example:
  - 1 4X Infiniband link/4 1X Infiniband links
  - 4 1000BX links
  - 1 10GbE XAUI link
  - 4 2Gpbs FC links
- Dedicated Management Dual Star 10/100/1000 Ethernet
- 10 pair Adjacent Slot Update bus
- Redundant blade-centric hardware management bus
- N+1 Redundant Power
- Cost focused module mechanical design

Cluster 2002

Futures

ATCA Target Serial I/O Standard Comparisons

Serial Standard	Data rate per Channel (Gbps)	Baud rate per Channel (Gbps)	Ref Clock (MHz)	Pairs/Channel Channel Configs
Infiniband	2.0	2.5	125	2 1X,2X,4X,12X
1Gb Ethernet 1000BASE-CX	1.0	1.25	62.5	2 1X
10Gb Ethernet XAUI	2.5	3.125	156.25	2 4X
Fibre Channel	0.85 / 1.7	1.06 / 2.12	53.125 / 106.25	2 1X
Serial ATA	1.2	1.5	75	2 1X
Serial RapidIO	2.5	3.125	156.25	2 1X,4X
PCI Express	2.0	2.5	125	2 1X-32X


hp
invent