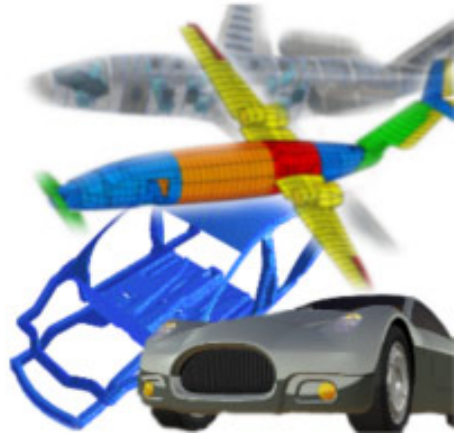
Tuning tips for Optimal NX Nastran performance

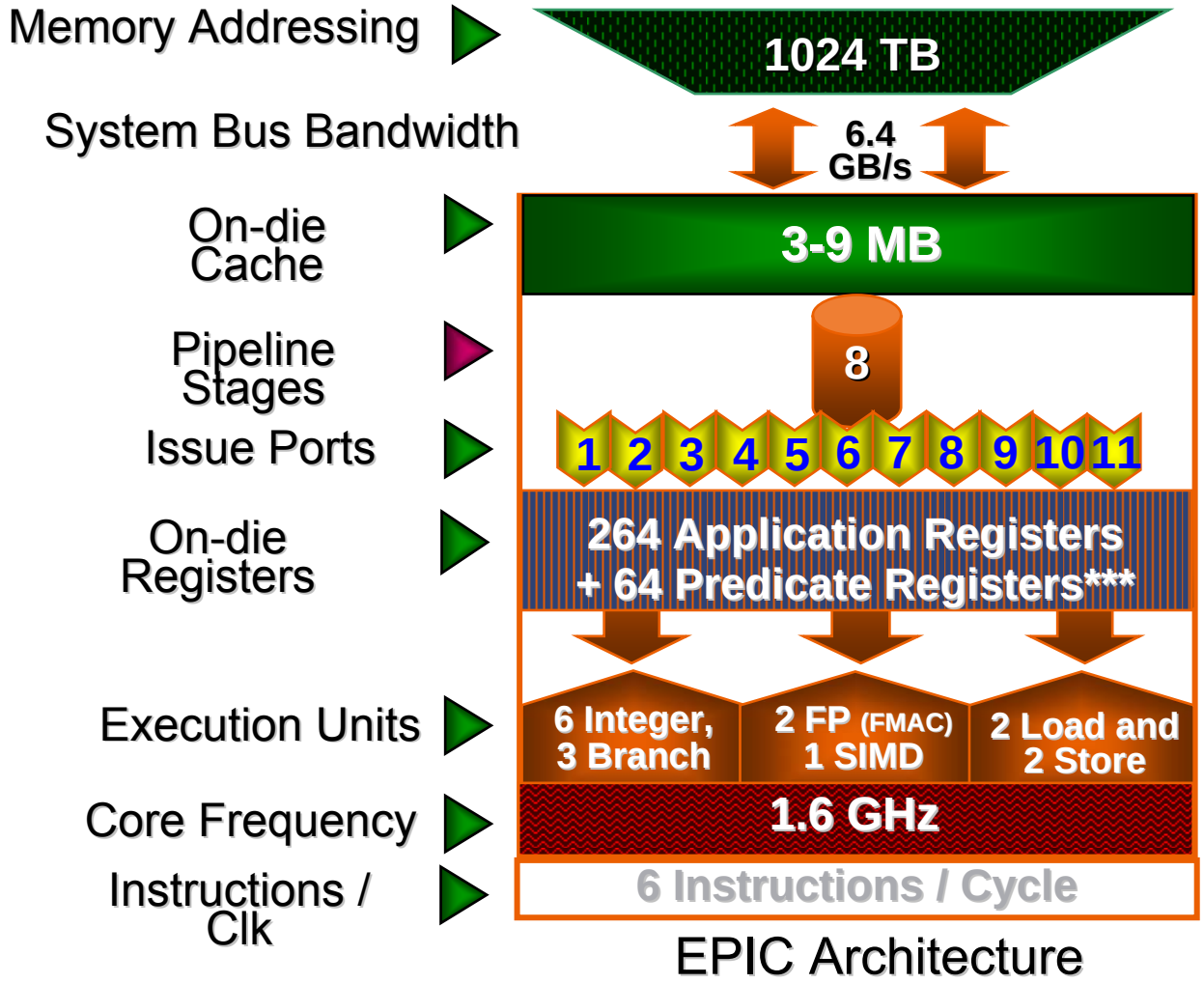
Pari Rajaram
High Performance Computing
Hewlett-Packard

NX Nastran

- Floating point intensive
  - cpu
  - memory system
- I/O intensive
  - Filesystem, controllers and disks.
- Distributed Memory Parallel (MPI)
  - Cluster
  - Large SMP servers



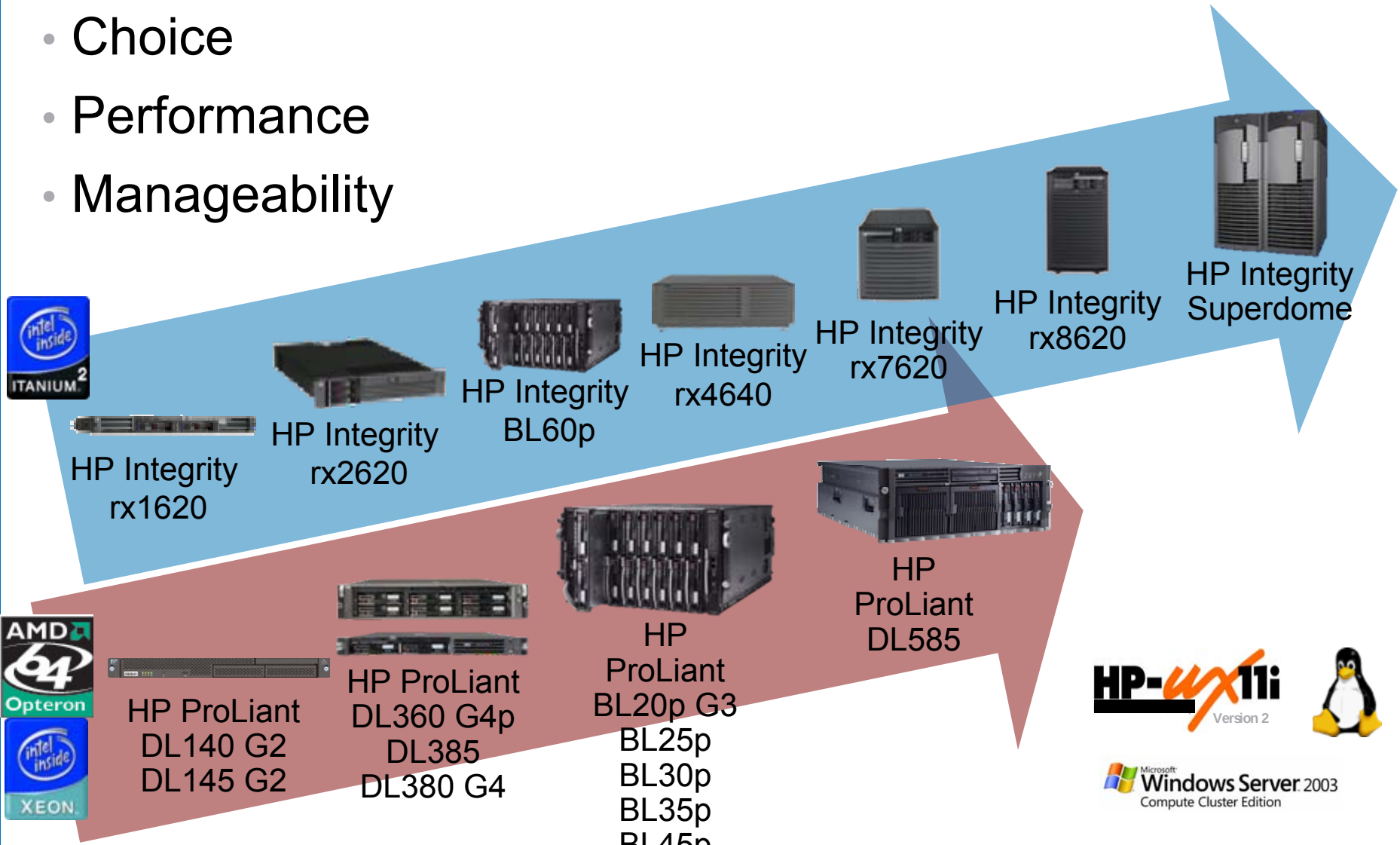
Intel Itanium® 2 Processor



*** Intel's EPIC technology includes 64 single-bit predicate registers to accelerate loop unrolling and branch intensive code execution.

HPC platforms

- Choice
- Performance
- Manageability



Nastran job's memory requirement

- **`nastran mem=estimate`**
- After the job completion, look in the f04 file and see if the memory allocation was enough.

MINIMUM MEMORY REQUIREMENT = 6275 K WORDS MEMORY AVAILABLE = 108480 K WORDS

MEMORY REQR'D TO AVOID SPILL = 62675 K WORDS MEMORY USED BY BEND = 12235 K WORDS

HIWATER 123620969 WORDS

- Do not give more memory than required to run in-core.
- Give some of the remaining server memory for OS buffer cache or to nastran smem.
- Can allocate more than 8GB memory now using ILP64 option.

Nastran Scratch files

- **nastran mem=estimate**

```
Estimated DOF=1584650
Estimated memory=730.3MB
Estimated disk=26591.2MB
```

- after the job completion, look in the <job>.f04 file to see the actual file transfer and size

FILE	ALLOCATED (BLOCKS)	HIWATER	HIWATER (MB)	I/O TRANSFERRED (GB)
SCRATCH	250000	62253	15563.250	39.330
SCR300	250000	117099	29274.750	2311.260
			TOTAL:	2350.771

Bring Elapsed Time closer to CPU Time

I/O bound

case1: Scratch file I/O to a /tmp single disk

real 3650 minutes:seconds.millisecond
user 2438
sys 320
I/O time = 3650 – (2438+320) = 892 seconds

case2: Scratch file I/O to a Raid disk (MSA30 7disks/1controller)

real 2895
user 2412
sys 286
I/O time = 2895 – (2412+286) = 197 seconds.

General guidelines for RAID configuration

For good application performance:

- Large datasets
  - Every nastran job needs 4-5 disks per controller of I/O bandwidth.
  - Nastran throughput jobs: multiplies all this !
  - HP MSA20 - SATA disks JBOD 150MB/second

  - HP MSA30 - Ultra 320 SCSI disks Raid 320MB/sec

HP Scalable File System (SFS) storage

Business Need

- ✓ Balanced I/O bandwidth for computationally challenging science and engineering

Storage Services

- ✓ Scalable bandwidth for technical computing

Enabling Technology

- ✓ Distributed bandwidth
- ✓ Standard building blocks
- ✓ Single system image for 1000s of clients
- ✓ Pay as you grow scalability
- ✓ Open Source technology

OS Buffer cache

- Buffer Cache (I/O)
  - Cache recently read/written I/O data for reuse in memory.
  - Used to hold ‘read-ahead’ and ‘write-behind’ data to increase i/o rates seen by an application.
  - Backup RAID I/O with decent amount of memory for buffer cache.
  - Leave room for Operating system and application.

Tuning OS Buffer Cache

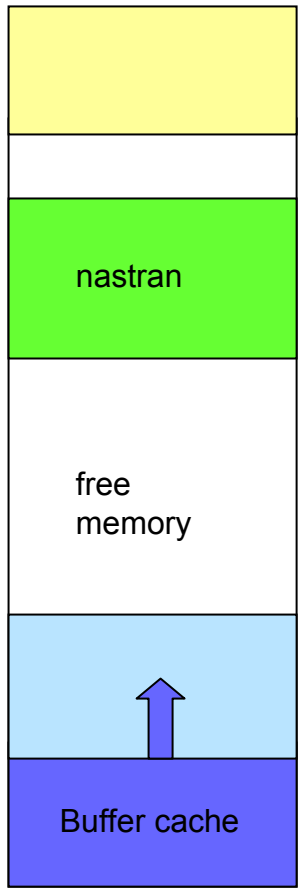
hpux Tunables

OS kernel

Nastran Memory

Buffer cache maximum = dbc_max_pct

Buffer cache minimum = dbc_min_pct



Physical Memory in the server

Filesystem tuning

- Natran I/O pattern
  - More reads than write request
  - Majority sequential
  - Strided
  - Random

- /etc/vx/tunefstab file
  - Read_ahead = read_pref_io*read_nstream
    - read_pref_io = 256KB
    - read_nstream = number of disks per controller
  - Write_behind = write_pref_io * write_nstream
    - write_pref_io = 256KB
    - write_nstream = number of disks per controller

NX Nastran

+ HP MPI
  - support for wide array of Interconnects on HPUX and Linux

  - production quality MPI that scales 1000+ nodes.
  - support from HP

+ HP Math library (MLIB)
  - Support for HPUX and Linux
  - 8 byte integer

HP: Platform of Choice for CAE™

www.hp.com/go/cae