



A QUEST FOR APPLICATION PERFORMANCE



Norman Lastovica
lastovica@sciinc.com
www.sciinc.com

My Background

- ◆ Pre-Digital
 - ◆ AMD 2901 bit-slice microcode
 - ◆ 68000 assembler
- ◆ Digital Equipment
 - ◆ Performance
 - ◆ VAX to Alpha
- ◆ Oracle
 - ◆ Performance
 - ◆ Alpha to I64
- ◆ SCI
 - ◆ Performance
 - ◆ VAX, Alpha, I64 to x86



Interact

Interview

Follow Evidence

Stories of the Road



Detect



- ◆ What is “it”?
- ◆ Did “it” ever work?
- ◆ What changed?
- ◆ Is “it” slow for all or just one?
- ◆ How slow is slow?



- ◆ Identify
- ◆ Quantify
- ◆ Document
- ◆ Graph



**THINGS ARE NOT
ALWAYS AS THEY
SEEM**

WALL CLOCK TIME ALONE ISN'T HELPFUL TO US



CPU TIME, ELAPSED TIME,
DIO, BIO, LATENCY, RATE



Where Computers Spend Time

- ◆ Executing useful instructions
- ◆ Waiting...
 - ◆ Disk I/O
 - ◆ Network I/O
 - ◆ User I/O
 - ◆ CPU
- ◆ “Overhead”



A person in a dark robe is performing a ritual. They are holding a wooden bowl filled with white smoke or incense. Several lit candles are visible in the background and foreground, creating a warm, atmospheric glow. The scene is dimly lit, with the primary light sources being the candles and the smoke itself.

Faster hardware is a bad first solution
to slow software

Modern hardware
is very fast

Hardware does not compensate
for everything



Bill Wright
Oracle Corporation

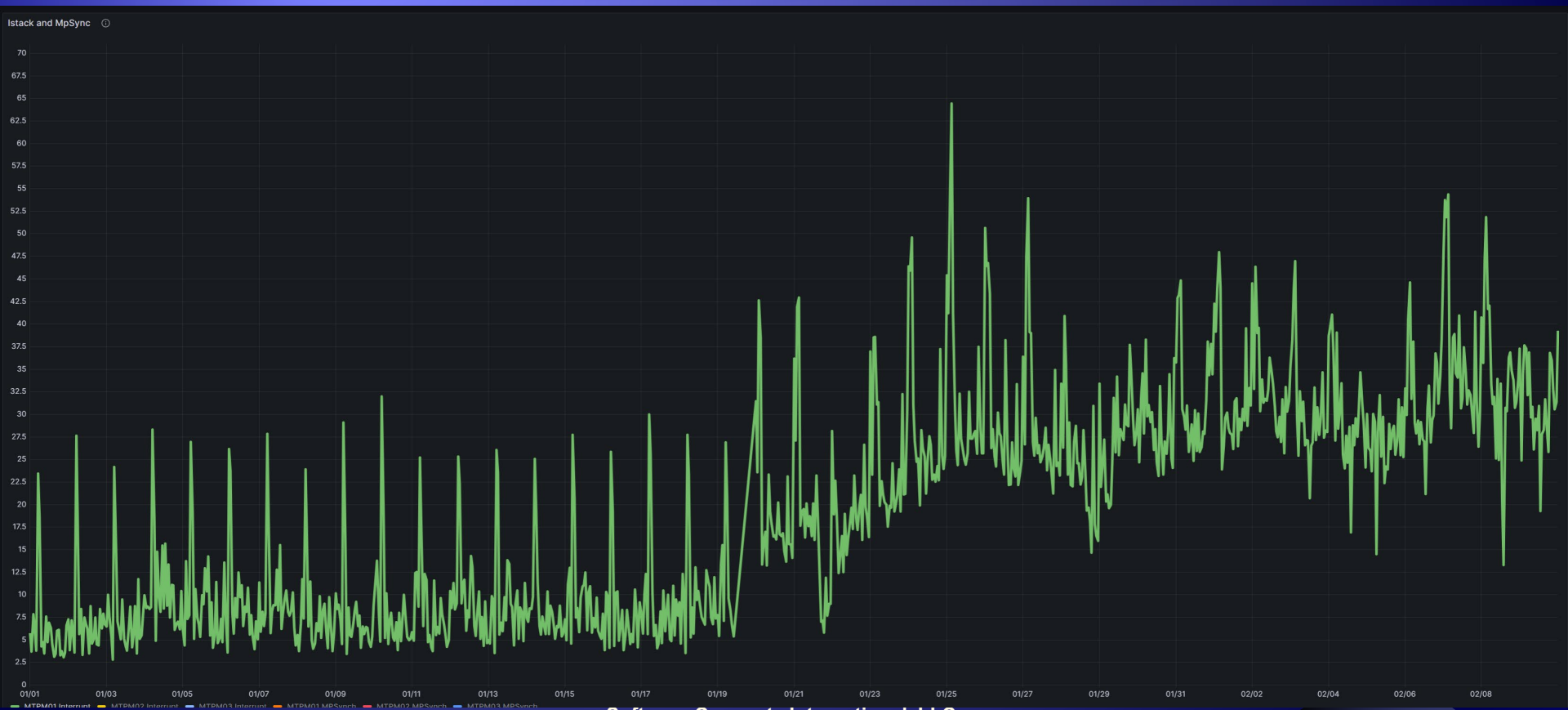
“It takes more than foul language
to improve performance”

Curious Case Of IMB



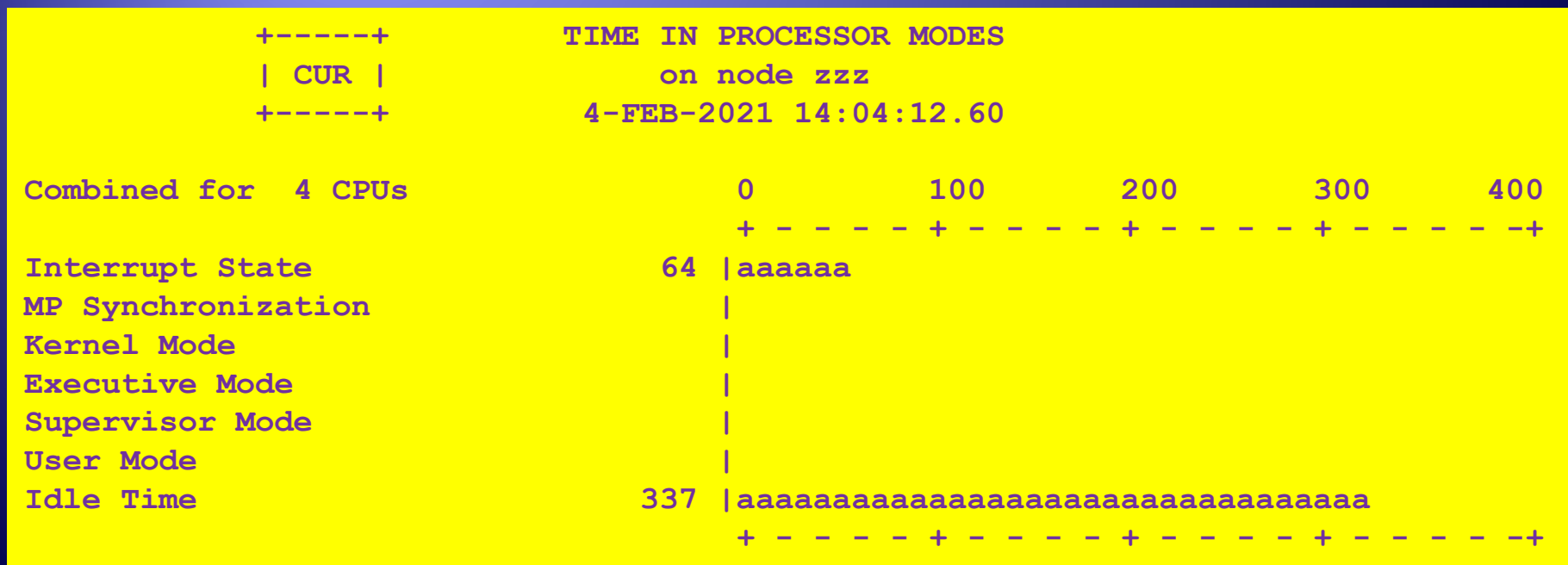
Interrupt Stack

- ◆ Migrated Alpha hardware to emulation
- ◆ Much higher Istack time



Reproduced In-house

- ◆ SCI Synthetic lock traffic test
- ◆ Charon emulator
- ◆ OpenVMS V7.3-2



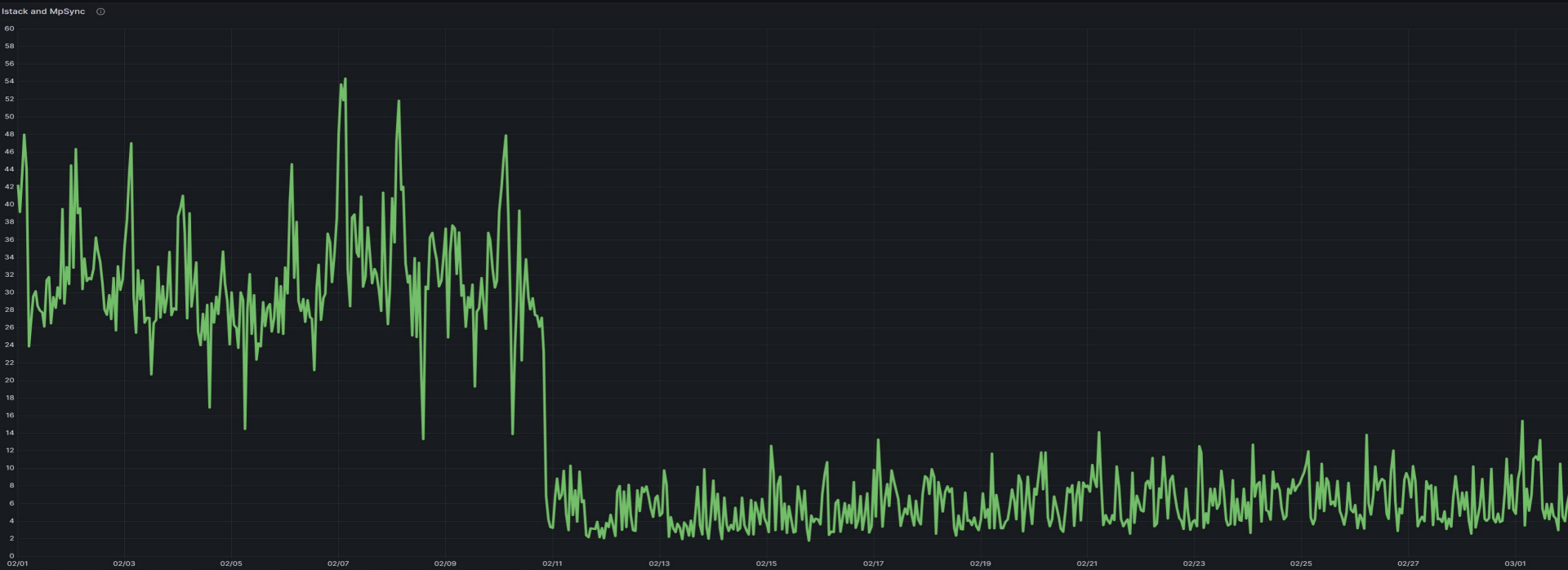
Why?

- ◆ Vast bulk of PC samples in SYS\$CLUSTER
- ◆ IMB (instruction memory barrier)
 - ◆ “Software that creates or modifies the instruction stream must execute a CALL_PAL IMB before trying to execute the new instructions.”
 - ◆ Workaround for flaw in specific Alpha hardware

```
;          X-36      xxx0792          xxxx x. xxxxxxx          19-Sep-2003
;          Fix EV6 bullcrap.  Workaround hardware issue until
;          firmware update can be supplied.  This introduces
;          an artificial stall after attempts to acquire the
;          LCKMGR spinlock NOSPIN, to allow EV6 hardware to
;          insure that correct behavior will ensue.  No affect
;          on other processors but for the annoyance of the
;          slowdown in the send and receive message paths.
```

Instant Relief

- ◆ SCI “cluster magic” patch
- ◆ >3X reduction in CPU on internal test system



On an emulator? See me after the presentation

FTP Slow

Performance Issues

- ◆ “...*seeing performance issues starting yesterday*”
 - ◆ Actually related to FTP throughput
- ◆ Development “fast”; production “slow”
- ◆ Checked *everything*
 - ◆ Disk/network latency & bandwidth
 - ◆ CPU/memory latency & bandwidth
 - ◆ LAN switches, cables, NICs
- ◆ Even rebooted 6 node production cluster to no avail



FTP Efficiency

- ◆ TCPIP\$FTP_WNDSIZ
 - ◆ *“Changing the window size of the send and receive buffers can improve network performance.”*
- ◆ Development system defined to 1000000;
production at 4096
 - ◆ Since years
- ◆ Nothing “started yesterday”
 - ◆ Beware “group think”



Alpha to I64

Financial Institution

- ◆ “Bet your economy” system
- ◆ Migration Alpha to I64
 - ◆ 99% BASIC → port



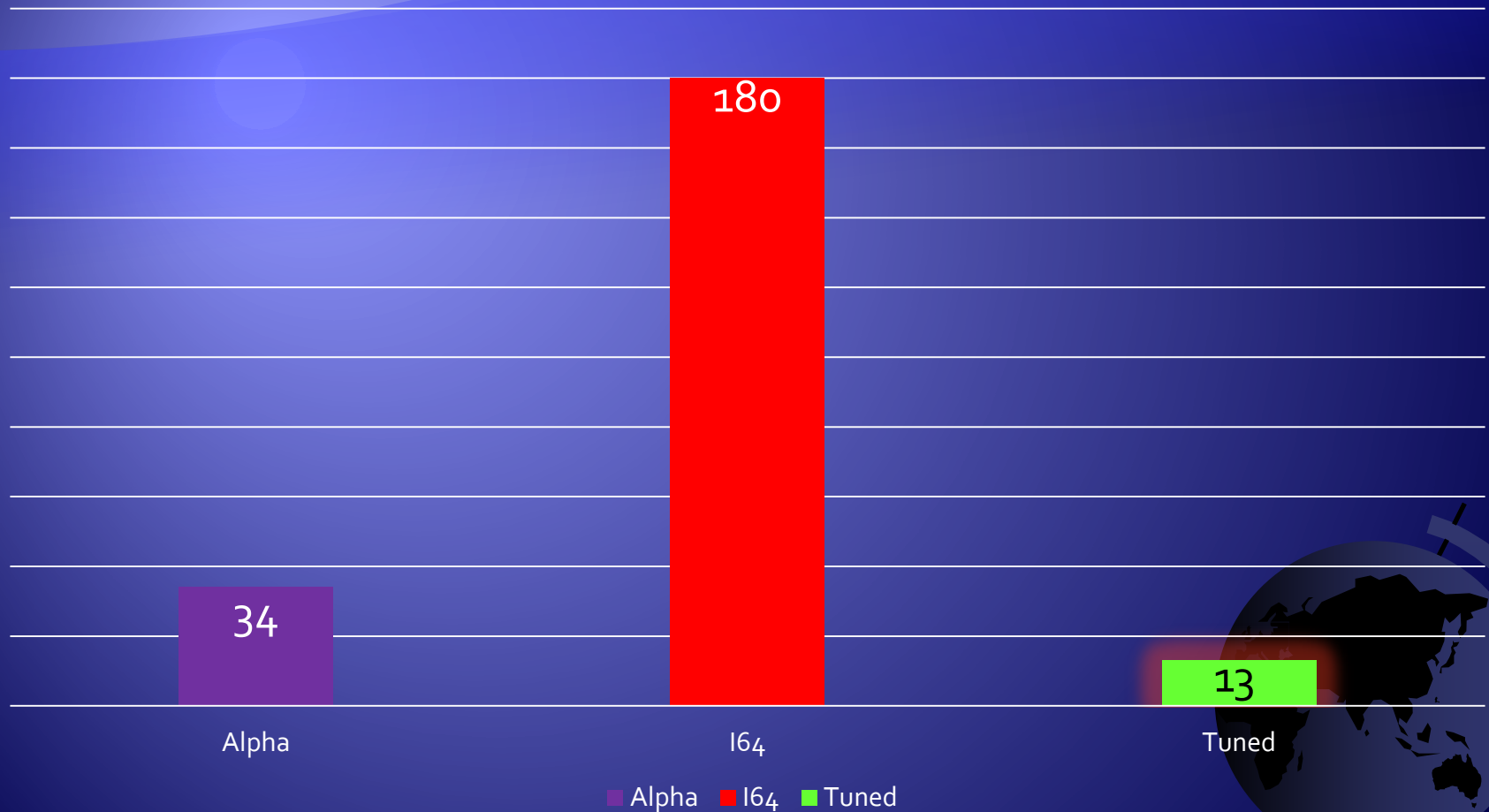
BASIC on I64

- ◆ Application much slower on I64 than Alpha
- ◆ Related BASIC, BASRTL, & I64 condition handling
- ◆ Corrected with minor application modifications
 - ◆ DBASIC\$IO_NO_SIGNAL in high frequency areas
 - ◆ USEROPEN to access files with NQL
 - ◆ Always a good idea



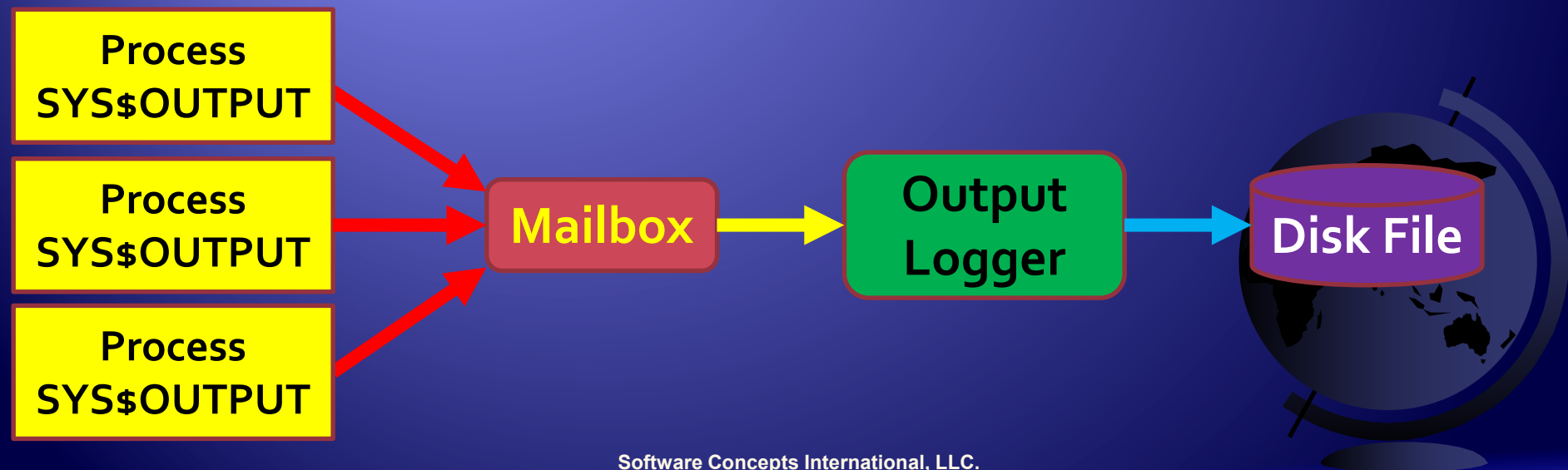
Results of Tuning BASIC

Nightly Processing Run Time Minutes



SCI output logger

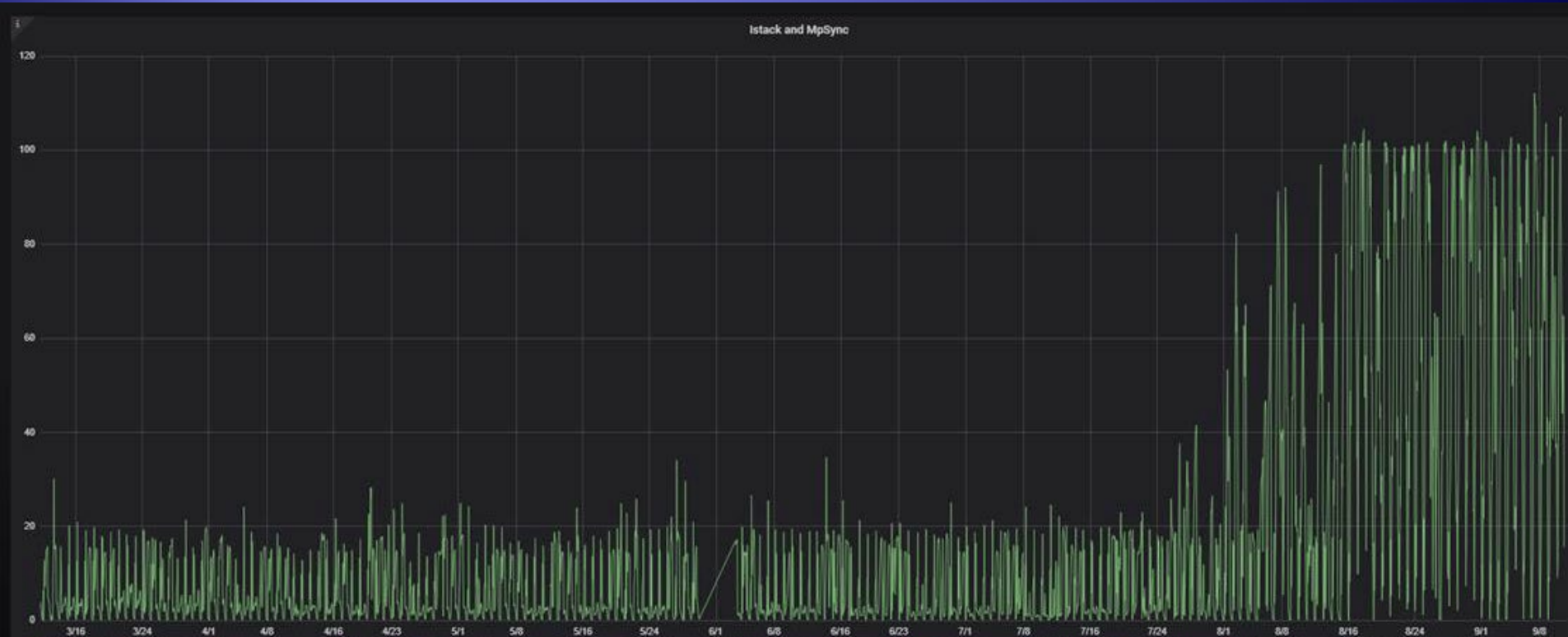
- ◆ Tens of thousands TCPIP connects per day
 - ◆ Application design results in log file for each
 - ◆ Frequently exceed creating 32k small files per day
 - ◆ Disk and directory contention and size
- ◆ SCI's output logger solution



Terminal Emulator

Interrupt State

- ♦ Terminal emulators deployed using TimingMark as keepalive message...every zero seconds
 - ♦ Corrected to send every 5 minutes



Slow RMAN



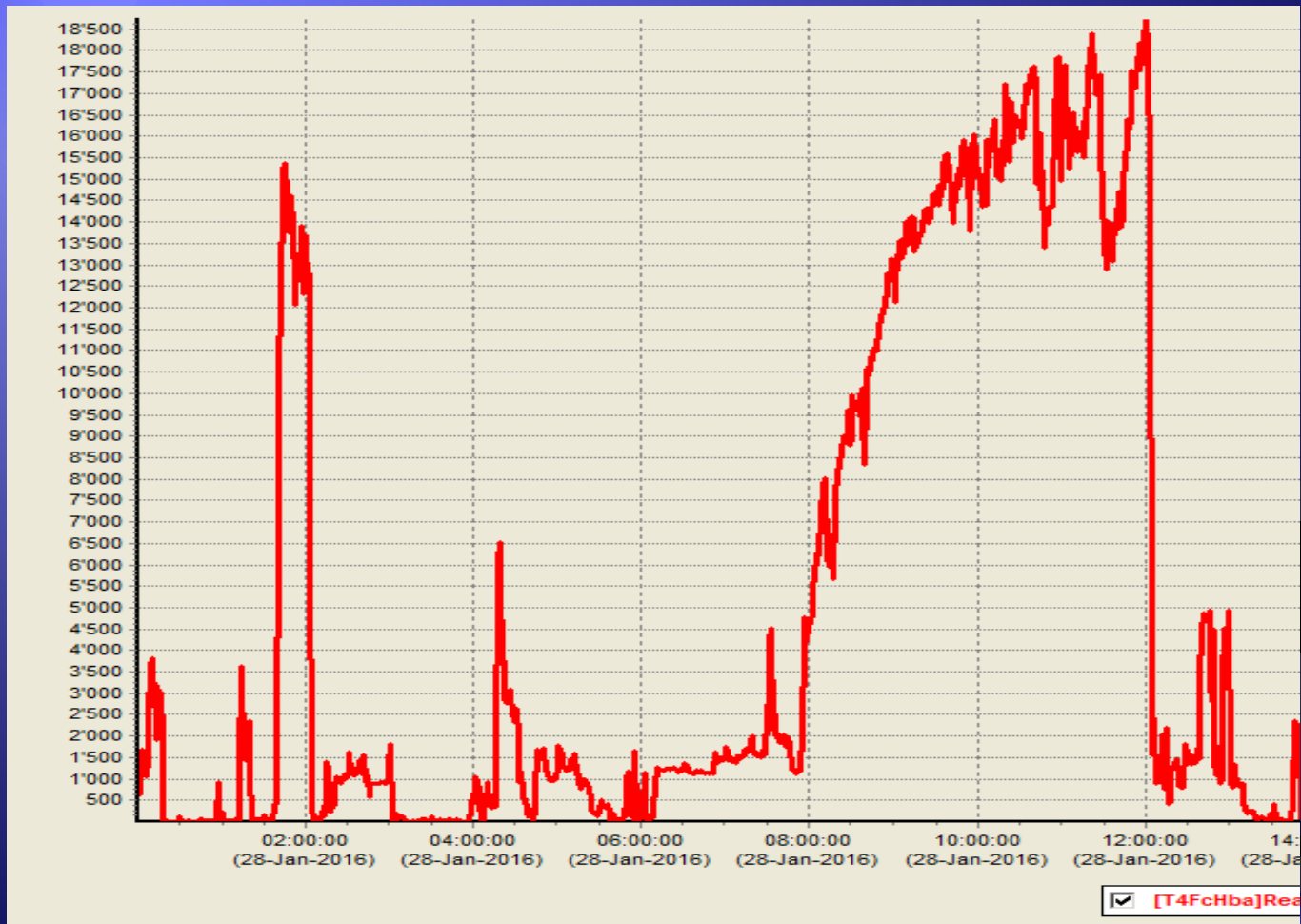
Oracle Database Backup

- ◆ Customer reports large Oracle database backup requires >24 hours
- ◆ SCI requested to assist



RMAN

- ◆ RMAN file extending results in massive amounts of file header reads from disk when FCP caches fill

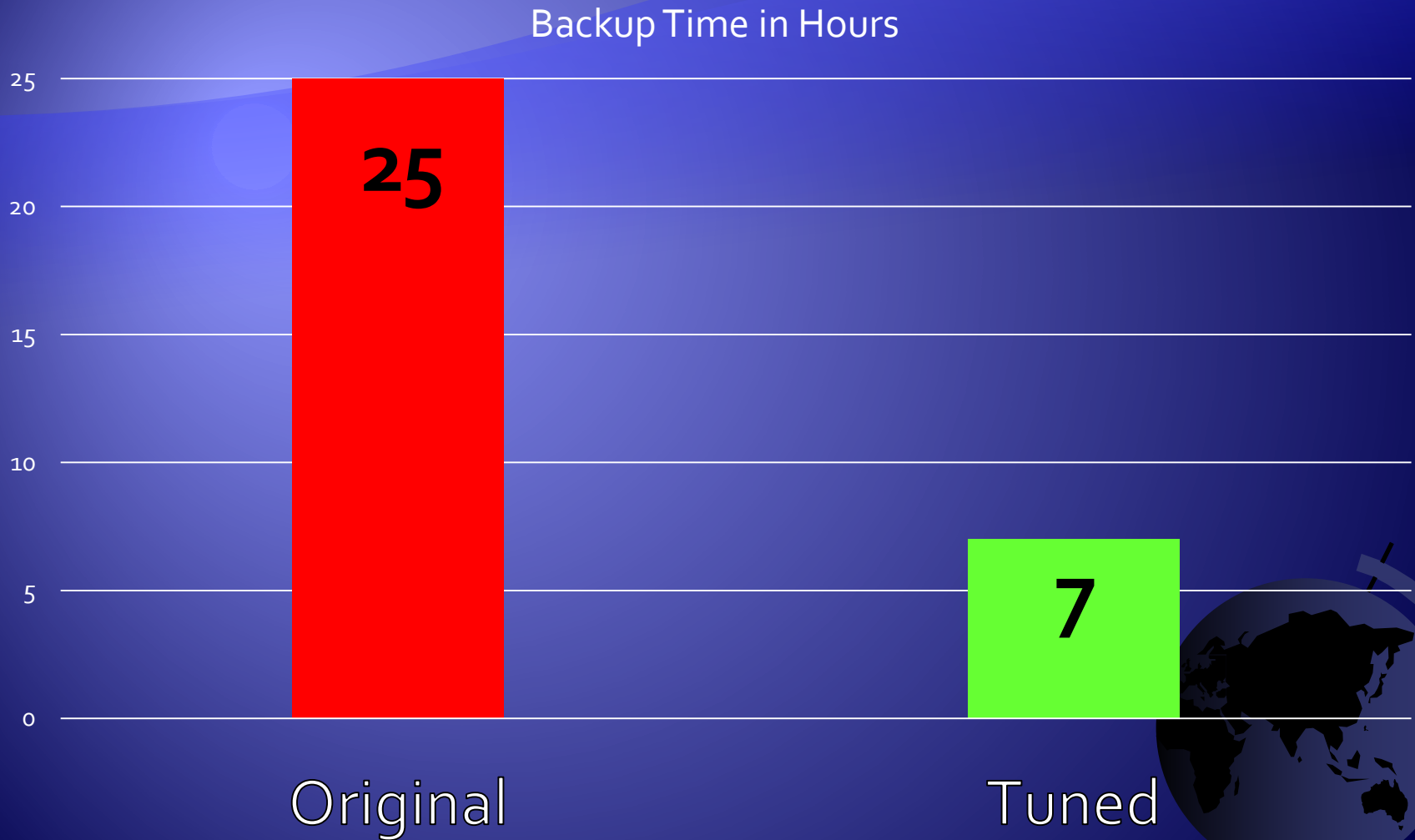


Degrades

- ◆ RMAN starts fast and degrades as it continues
- ◆ Oracle closes output file, extends it an entire 1024 blocks (512KB at-a-time), then reopens, VMS has to scan extents to find EOF
- ◆ Once XQP header cache fills, positioning to end-of-file consumes bulk of time
- ◆ Reinitialize/remount volume with appropriate settings



>3x FASTER With SCI Tuning





Some Truth About Duplicates Indexes

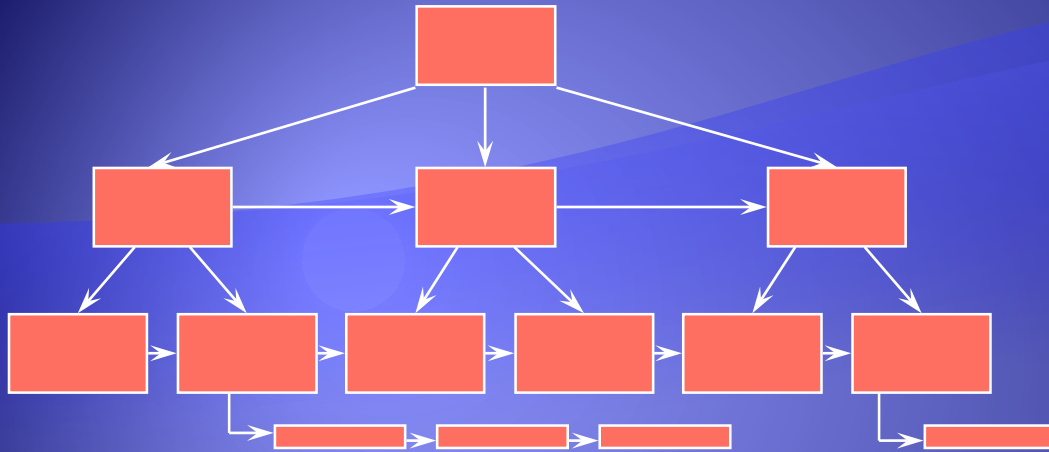
“TODO” Flag

- ◆ Application utilizes “PROCESSED” column
 - ◆ F = Needs to be processed
 - ◆ Just a few rows
 - ◆ T = Complete
 - ◆ Zillions of duplicates
- ◆ SCI reorganizes index SORTED RANKED
 - ◆ Performance goes to heck



Rdb SORTED Insert

Constant insert speed



New SORTED duplicate entries in a chain inserted here

- ◆ "Reverse chronological order"
- ◆ Duplicate entries store DBKEY in LIFO chain
- ◆ Insert or update duplicate node at front of duplicate node chain

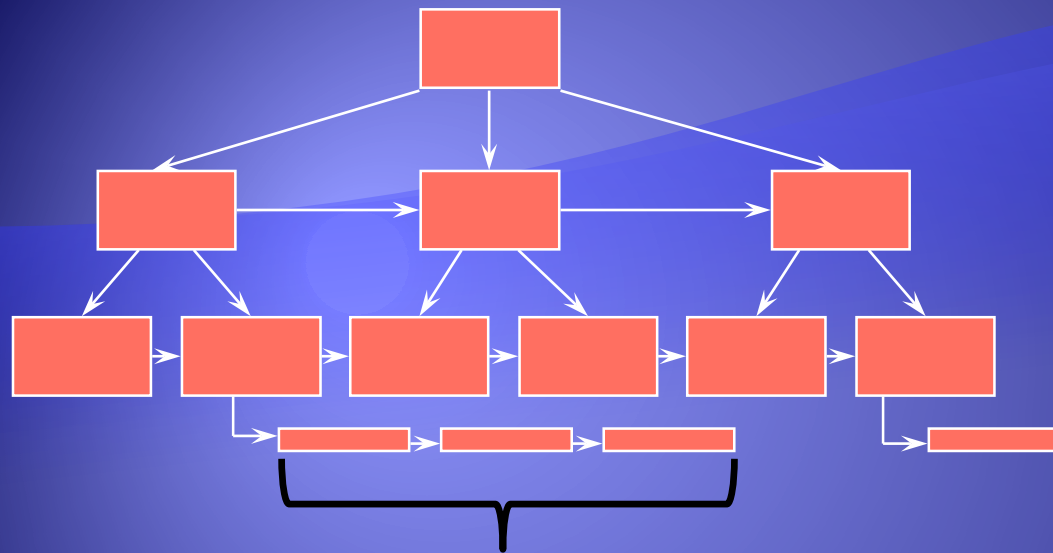




New RMS Indexed
duplicate entries in a
chain inserted here

- 
- ries store RFA
- end for insert

Rdb SORTED RANKED Insert



Variable insert speed;
degrades as table grows

New SORTED RANKED
duplicate entries inserted in
correct location via scanning list

- ◆ "Physical storage order"
- ◆ Duplicate entries stored in DBKEY order
- ◆ Primary node stores pointers till full then to chain
- ◆ Scan BBC list (duplicate node chain) to locate insert point

Securities Trading vs Speed of Light



Exchange

- ◆ Cross-site cluster; cross-site shadowing
 - ◆ Mini-merge bitmap updates make things worse
- ◆ Move DR site to 'slightly' more distant data center
- ◆ End of day processing much slower
- ◆ Trade volume greatly exacerbated by COVID



Single Block Synchronous IO

- ◆ Application log file accessed with single block synchronous IO (e.g. “transaction log”) routines
- ◆ Nightly processing step more than twice as long
- ◆ Dramatically impacts end-of-day processing
 - ◆ Dangerously close to window of downtime



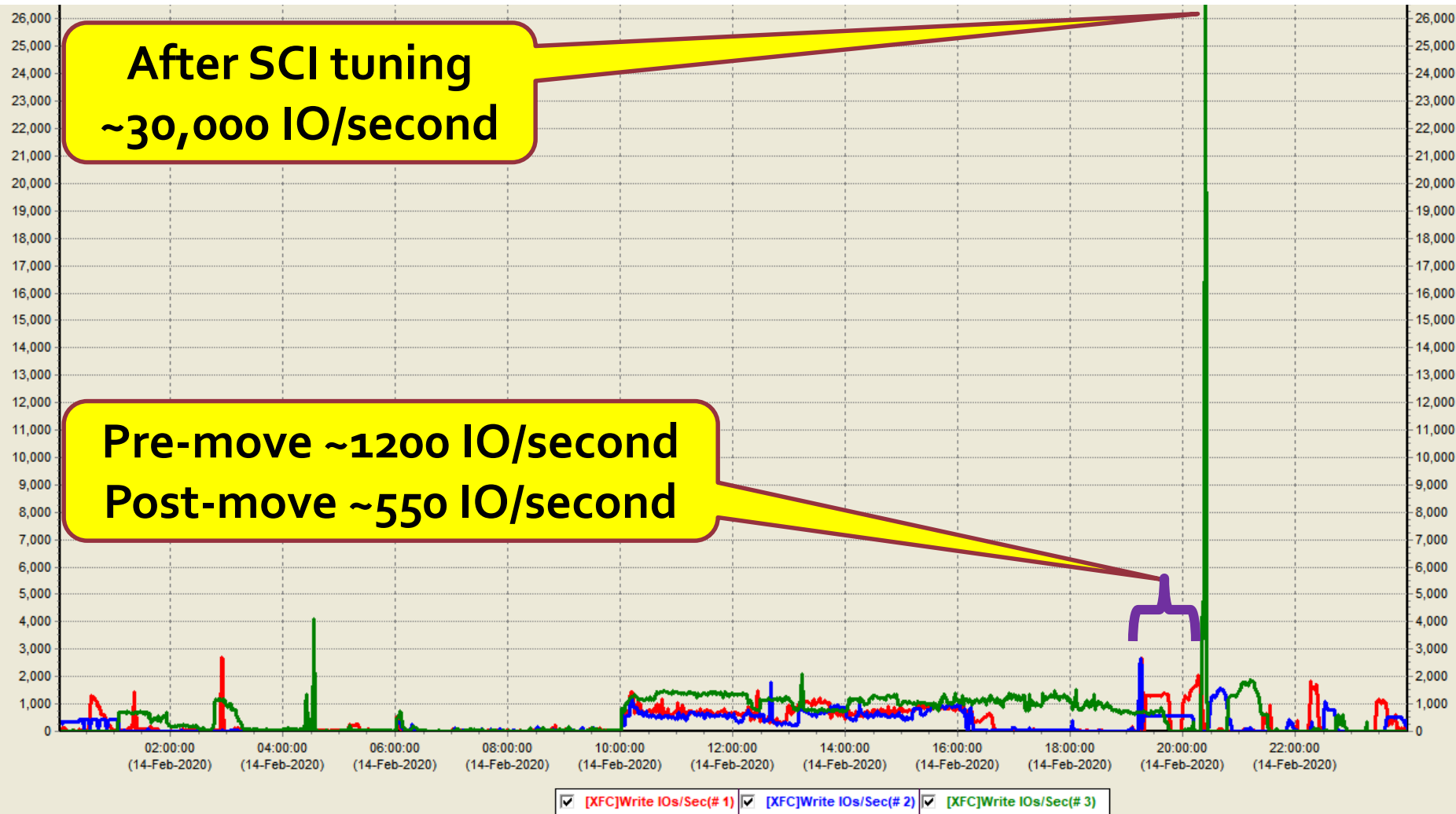
Solution

- ◆ Reimagine processing
- ◆ Single command procedure involved
 - a) Move transaction log to DECram device
 - b) Perform optimizations
 - c) Move transaction log back to media



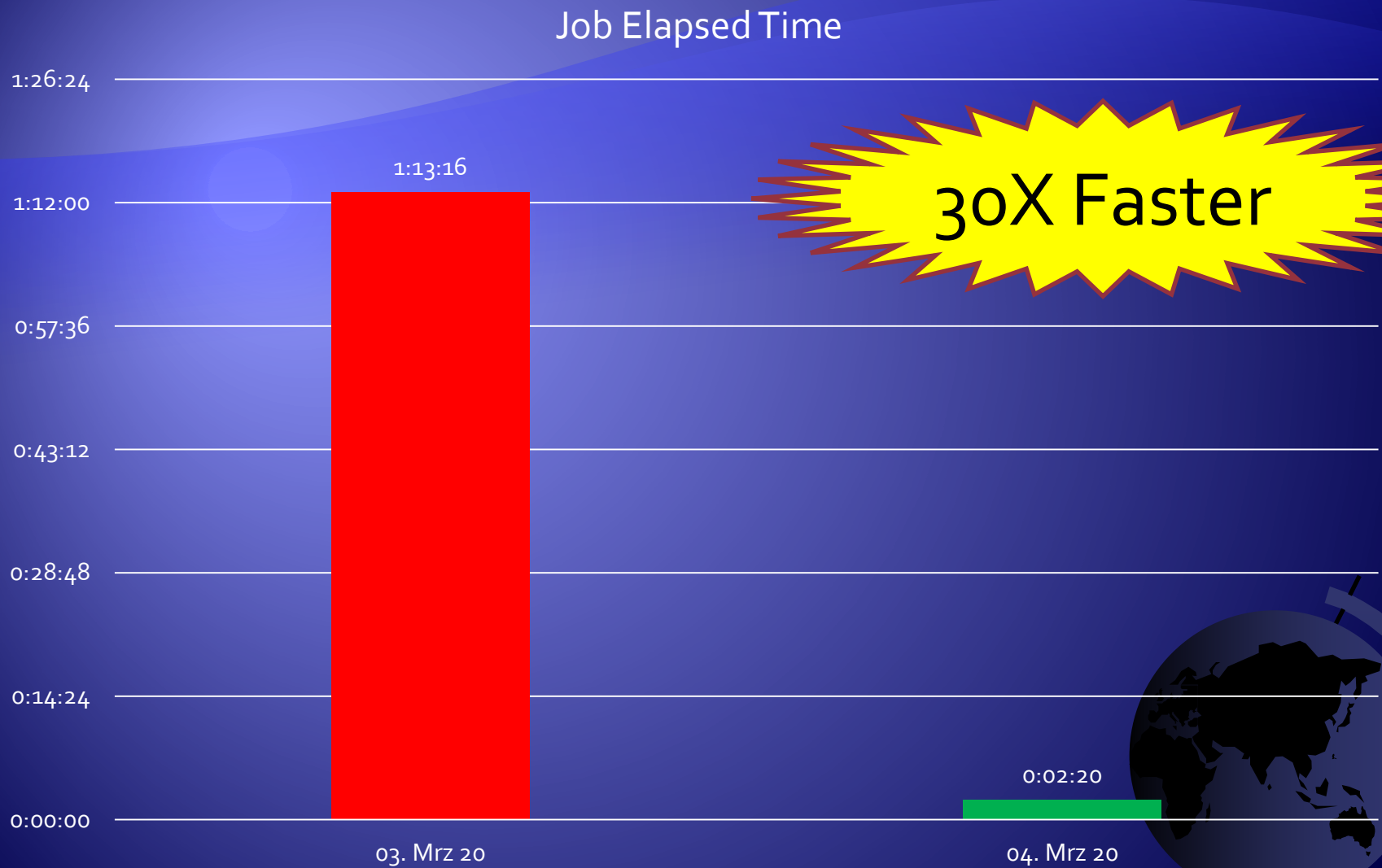
No longer limited DIO Rate

#1 14-FEB-2020 **#2 17-FEB-2020** **#3 10-MAR-2020**





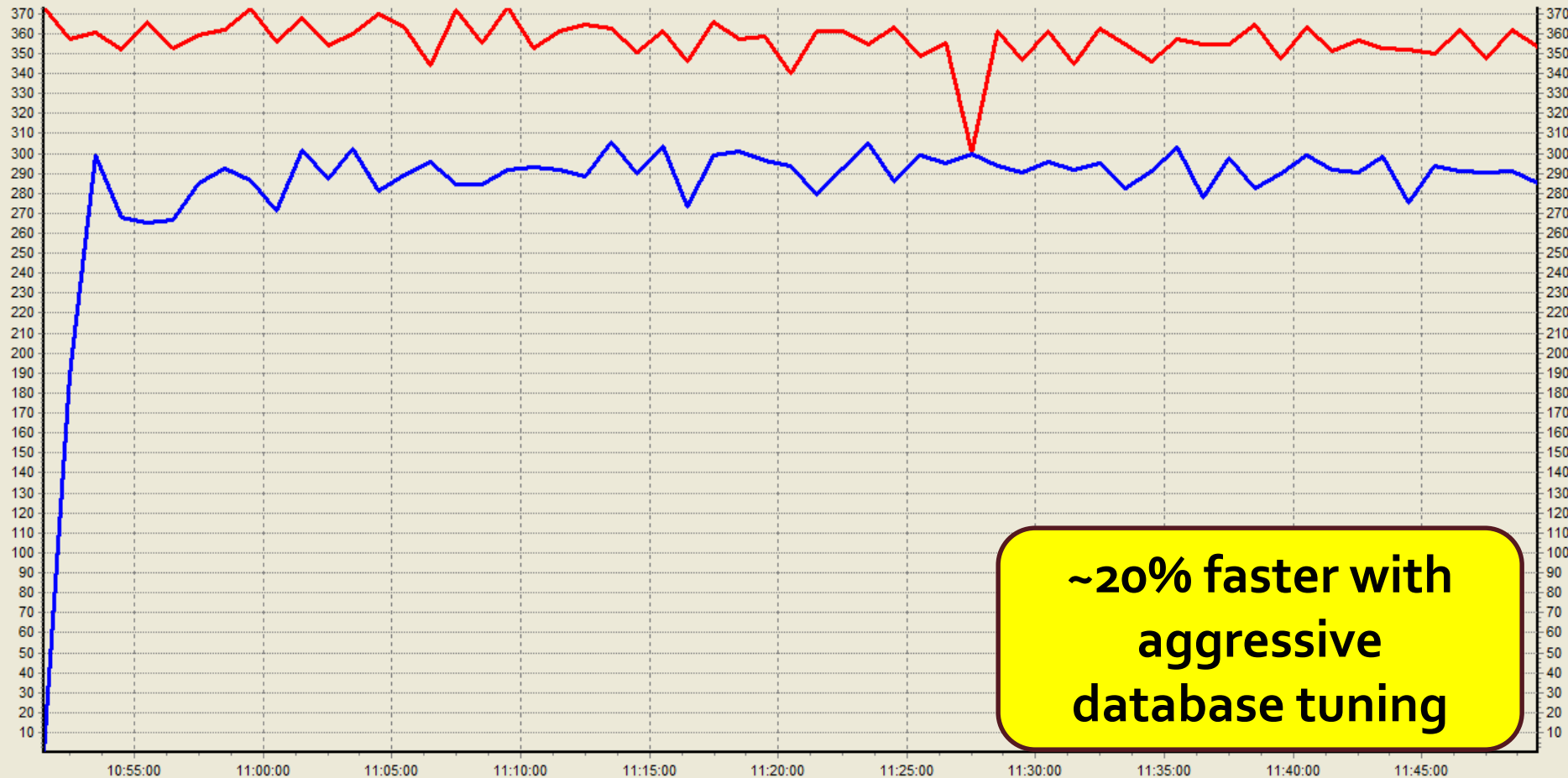
Gained > 1 hour



OLTP

Production Configuration TPS

Database / System Tuning



Trades Registration per Second Sample

Rdb V74 Slower Than Rdb V73?

- ◆ Expectation that Rdb 7.4 is faster or at least not slower than Rdb 7.3
- ◆ “Job XXX in TEST consistently takes ~5X longer than in PROD”
 - ◆ Elapsed 00:01:04 vs 00:00:12
 - ◆ CPU ~same, IO less
- ◆ Blocked on Rdb snapshot cursor lock
- ◆ Resolved via job scheduling realignments



Get To Work

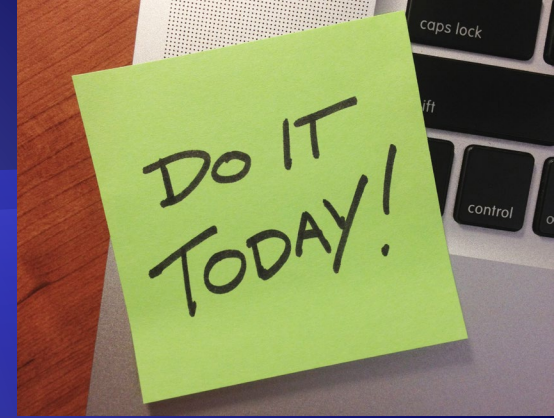
A close-up photograph of several hands stacked together in a circle, symbolizing teamwork and collaboration. The hands are of various skin tones and are wearing different colored sleeves: red, orange, brown, blue, and white. One hand has a ring on the ring finger. The background is dark and out of focus.

A ACTION
C CHANGES
T THINGS

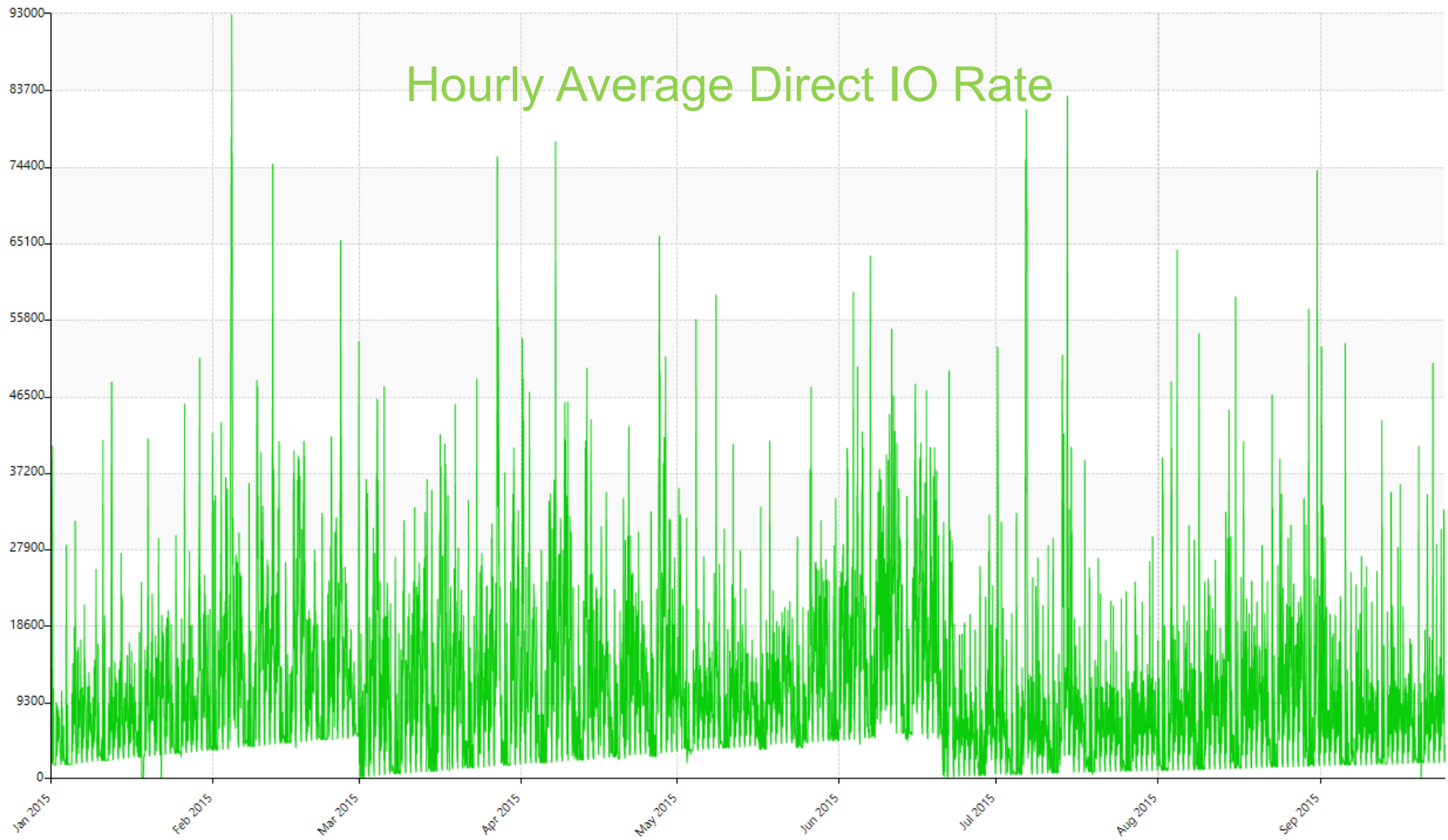
A hand holding a white marker, pointing at the text 'T THINGS'.

Gather Data - T4 / TlViz

- ◆ Required for system performance analysis / trending
- ◆ Capture statistics in database for trend analysis
- ◆ Application and database statistics as well

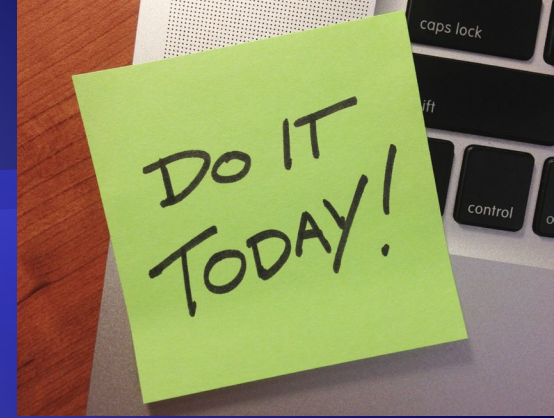


The Big Long Picture

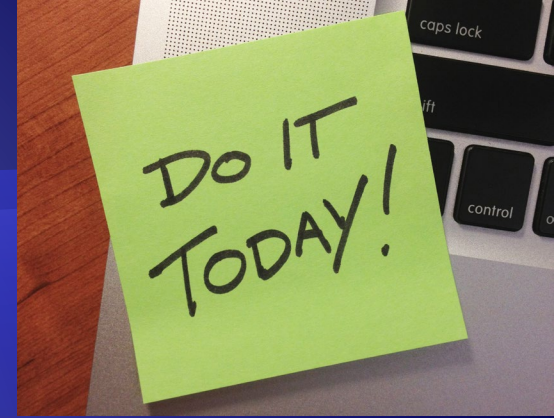


RMS Tuning Areas

- ◆ For indexed files consider automated
 - ◆ ANALYZE/RMS/FDL...
 - ◆ EDIT/FDL/NOINTERACTIVE...
 - ◆ CONVERT/FDL
- ◆ Use larger buffer sizes & counts
 - ◆ MBC, MBF
- ◆ Use DFW, WBH, RAH, DEQ
- ◆ Use NQL



Compiler Optimization

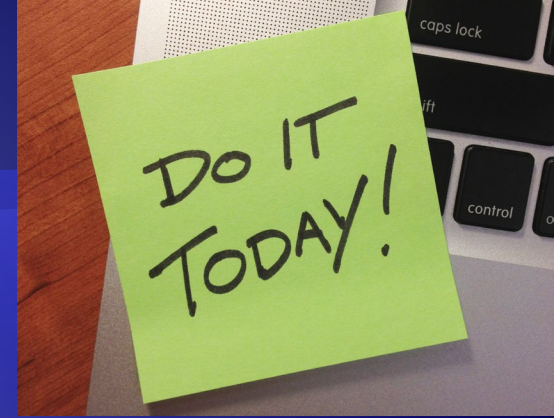


- ◆ Always compile
 - ◆ **/OPTIMIZE**
 - ◆ /ARCH= ! When on Alpha
 - ◆ /DEBUG
 - ◆ /LIST
 - ◆ /MACHINE
 - ◆ For actively debugging code, add /NOOPTIMIZE
- ◆ Link
 - ◆ /MAP /FULL /CROSS /SECTION_BINDING



Evaluate

- ◆ Compare CPU vs Elapsed time
- ◆ Big hints
 - ◆ CPU bound
 - ◆ IO bound
 - ◆ Something not easy to explain?
- ◆ Review execution strategies





Interview

Isolate

Identify

Questions? Thoughts?



Software
INTERNATIONAL
Concepts

Worldwide
Managed Services for
OpenVMS and Rdb

About Software Concepts International

Managing OpenVMS systems and
databases requiring the highest
levels of performance and
availability – worldwide



Software Concepts International

- ◆ Located in Nashua, NH (USA)
 - ◆ 30 years in business supporting OpenVMS!
- ◆ International reputation as leading provider of
 - ◆ Managed services for OpenVMS & databases
 - ◆ OpenVMS performance and consulting services
- ◆ Proven global track record
 - ◆ Actively managing 100s of systems and databases
 - ◆ At many sites 24x365 since 1995



Software Concepts International

- ◆ System and database performance consulting
- ◆ VAX/Alpha emulation – CHARON reseller
- ◆ Migration consulting
 - ◆ Specializing in minimal downtime migrations
- ◆ Oracle's worldwide provider of CODASYL DBMS training

