

Junk it: How BAE saves \$3.2m a year by throwing out excess data



Charlee Gothard

25 March 2015 • 6 min read

Active Navigation and an employee reward scheme was all it took

"Every IT person should be saving at least 10 per cent of their annual yearly salary," BAE Systems data centre storage director Mark Tango tells *Computing*. "And that is not a stretch. IT people waste more money, with their little science fair projects or purchasing the wrong equipment, having failed to do analysis or find the right provider. And they overlook all the pieces in play in their organisation," he says.

It's a scathing statement, but perhaps not untrue. One only needs to look at [the £100m digital media initiative failure pinned on the BBC's sacked CTO John Linwood \(who subsequently won his case against the BBC for unfair dismissal\)](#), or [any number of costly UK public sector failures](#).

For BAE Systems - a multinational defence, security and aerospace giant - it turned out that even storing its day-to-day data was sucking money out where none needed to be spent.

It's a more common problem than many IT chiefs realise. [Gartner recently identified that the cost per TB of storage in 2014 was \\$2,520](#) - a figure that covers hardware, software, personnel, connectivity, occupancy and disaster recovery.

Saving this cost or rather, in accounting terms, "avoiding" this cost became top of Tango's agenda when he joined BAE Systems in 2012.

"[After a] series of acquisitions of smaller defence companies around the US - which were all running independently, BAE Systems had started an initiative to build out two data centres, and started consolidating," explains Tango.

"But they brought junk over from these other sites, and the gentleman who was here before me was just going with the flow. He was a long-time employee with the 'This is the way we've always done it' mentality," laments Tango.

Getting senior managers together, Tango asked a key question:

"I said, 'Gentlemen, what is your formula for requesting, calculating or estimating what storage you'll purchase for the following year when budget and planning comes along?'. "

Their answer, he explains, was simply "We have no formula - we simply buy 40 per cent more storage every year".

After picking up his jaw, Tango spent the next nine months digging deep into BAE's data usage across the company.

"I spent time going through each device and looking at the data, and looking at how it was structured, what we had and didn't have," he tells *Computing*. Something needed to be done. Tango says junk data, particularly casually downloaded music and video files, infested a great many devices, so checking and deleting by hand obviously wasn't an option.

Luckily, BAE Systems had an existing relationship with file analysis firm Active Navigation, though on an admittedly small scale.

"BAE had hired their early product," says Tango, "for our legal department to identify data of people who'd left the company. The first iteration was pretty clunky, and it took many servers to get it running."

Nevertheless, potential was there.

"We quickly realized the platform (which had now been upgraded) let us rip everything out and junk it. Then we built three active scanning servers per data centre," explains Tango.

The goal was to reduce BAE Systems' enterprise data footprint from 5.5PB to 3PB, but on the way to meeting this aim, BAE Systems found itself becoming pretty seriously involved with Active Navigation's ongoing endeavours.

"We then began to get our feet wet with the new project, because we found some improvements that could be made to it, and [Active Navigation] released the revisions to all the customers," says Tango.

All guns were now blazing. The software was working a treat, and with BAE Systems' added involvement in even developing further versions, it was time for Tango to make perhaps his only real mistake.

"We got all excited and tried to boil the ocean," he laughs, explaining how IT "kicked off a scan of everything," resulting in a crunched and analysed 3PB of data, but "not enough people to review it".

With a 120-day cut-off point "before the data becomes stagnant," there was no choice but to start again - except this time with more people to review the harvested data.

With departmental managers put in charge of checking the "junk" data, Active Navigation had pulled out. But the real winner was dumping the data to an algorithmic Excel spreadsheet, instantly allowing managers to show the precise cost per GB for all the data stored at BAE Systems.

"We started this on our own with no guidance from upper management, but in the end when my boss got wind of it, upper management said 'this is a great project' and asked me to send names of people making progress every month," reveals Tango.

A reward scheme actually followed, with the best-performing data savers in given departments incentivised to keep up the good work.

There wasn't even any particular disadvantage in clearing away data which wasn't junk, but was over six years old, as Tango explains it was all placed into a specially-built "cold storage area" - some repurposed Dell EqualLogic servers -

which literally took only seconds longer to access than the truly local data. Tango decided not to place the data truly "off the grid" as he didn't want to be "inundated with phone calls" to try and access data that would take many hours to restore on a regular basis.

Tango states that the main reason Active Navigation's software was so attractive was due to the company's decision to engage multiple servers at once to check all the data.

"We're able to skim multiple storage devices all at once," says Tango.

"Once you buy the product you can spin up as many servers as you like. Other rival products make you buy another licence, buy another licence, buy another licence each time - but obviously the more servers you have up on line, the more you can save."

As a result, BAE Systems saw 881TB of junk data eliminated in 2014, amounting to a saving ("cost avoidance") of \$1.7m. Last year, this figure was \$3.2m in total cost avoidance from junk data.

A particular recurring cost area that Active Navigation identified was in databases that were terabytes in size, but often would only be using "500GB of data, and had been that way for years," comments Tango.

"When we switched to charging to what was allocated, everyone came running to get their databases rightsized."

Tango is now on something of a crusade to help the wider industry keep its data clean and its storage cheap. Apart from his opening statement about the amount IT people should save, he's joined the [Information Governance Initiative - an IT industry consortium and think tank](#) - and is now looking at ways to turn information management into governance, both for BAE Systems and enterprises in general. Within his company, Active Navigation will continue playing a key role.

"Governance around information today is paramount," he says.

"Right now there are no policies and procedures in place, but we need to expand that across the enterprise. Other people just let data proliferate, and that can't carry on."