Use of Business Intelligence Tools in the DSN

Joseph I. Statman and Silvino C. Zendejas[1]
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109, USA

ABSTRACT

JPL has operated the Deep Space Network (DSN) on behalf of NASA since the 1960's. Over the last two decades, the DSN budget has generally declined in real-year dollars while the aging assets required more attention, and the missions became more complex. As a result, the DSN budget has been increasingly consumed by Operations and Maintenance (O&M), significantly reducing the funding wedge available for technology investment and for enhancing the DSN capability and capacity. Responding to this budget squeeze, the DSN launched an effort to improve the cost-efficiency of the O&M. In this paper we:

- *Elaborate on the methodology adopted to understand "where the time and money are used". Surprisingly, most of the data required for metrics development was readily available in existing databases. We have used commercial Business Intelligence (BI) tools to mine the databases and automatically extract the metrics (including trends) and distribute them weekly to interested parties.*

- *Describe the DSN-specific effort to convert the intuitive understanding of "where the time is spent" into meaningful and actionable metrics that quantify use of resources, highlight candidate areas of improvement, and establish trends.*

- *Discuss the use of the BI-derived metrics. One of the most fascinating processes was the dramatic improvement in some areas of operations when the metrics were shared with the operators. The visibility of the metrics, and a self-induced competition, caused almost immediate improvement in some areas.*

While the near-term use of the metrics is to quantify the processes and track the improvement, these techniques will be just as useful in monitoring the process, e.g. as an input to a lean-six-sigma process.

I. Introduction

At the core of Business Intelligence (BI) tools, one finds the old adage that "one person's junk is someone else's gold". BI tools take advantage of the fact that modern enterprises, powered by computer systems, continuously create huge amounts of persistent data that is rarely used. This data often can be mined for precious information about underlying processes and trends, and exploited to improve the efficiency of any operations – space operations included.

BI software allows the analyst to create automated tools (or queries) that will:

1. Extract specific data
2. Perform desired analysis on the data (e.g. histograms, cross references)
3. Create reports, presenting the analysis results in tabulated and graphical forms
4. Distribute the reports to the users, on a regular basis, or when a threshold is crossed

[1] The work reported in this paper was conducted at the Jet Propulsion Laboratory, California Institute of Technology (JPL/CIT) under contract with the National Aeronautics and Space Administration and at NASA HQ.

The power of BI tools is that the data is generally available in the databases (or "data stores") already created, collected, and stored by the enterprise. The BI tools simply peruse these data stores to collect useful information.

Developing an equivalent set of tools from the ground up is a very expensive proposition. It makes sense to leverage the solutions provided by industry in this area. There are multiple companies that produce BI tools. Most of them directly, or through licensees, will support the delivery of the specific queries, allowing users to either develop an in-house analysis capability.

II. The DSN Data Stores

The Deep Space Network (DSN) is operated by the Jet Propulsion Laboratory (JPL) on behalf of the National Aeronautics and Space Administration (NASA). The DSN operates three antenna complexes located near Goldstone, California, Canberra, Australia, and Madrid Spain. Each complex provides mission support, relying on a 70m antenna and multiple 34m antennas. The DSN also has several facilities at and near JPL to support the operations.

The DSN uses a distributed Monitor and Control (M&C) system that includes hardware, software and staffing at the three antenna complexes as well as at JPL. During the operations, huge amounts of M&C data are stored (much larger amounts of data are normally generated but discarded).

Early in 2008, the DSN embarked on an effort to improve the efficiency of real-time operations. The objective was to translate prior efforts at automation into a significant reduction in the cost of the 24x7 staff. To make this effort successful, it was imperative that the work load on the staff be quantified, and a realistic assessment be obtained on the feasibility and risk of increasing the responsibility and associated workload of each staff member.

In searching for methods to get the quantification, we have leveraged on the experience of a pilot project that used BI tools to create monthly management metrics. This effort created the needed infrastructure – both the technical expertise and the contractual tools for quick generation of queries.

Finally, we found that the use of BI queries is very suitable to an iterative and incremental approach: rather than write requirements and develop the full set of queries, we established a set of (independent) deliveries. We used the guidance from the user community for inputs on the usefulness of developed queries and desired new and modified queries. In this paper we will use examples from Deliveries D1, D2, and D3.

III. Delivery 1 - Basic Use of M&C Database

The objective of delivery 1 was to quickly quantify the load on the real-time operators. A quick review of the M&C data store contents revealed that it includes every input from the operator (typically, a directive). Furthermore, within these operator inputs it was straight forward to identify interesting classes of directives, in particular the directives associated with activate automation scripts as well as those manually issued by an operator

Because report generation with BI tools requires minimal effort, there is a temptation to generate large number of reports and inadvertently overwhelm the user. We made an early decision to average data over either one week or four weeks and "push" data to the users on a weekly basis via e-mail. Based on our experience, we now use mostly 4-week averaging periods.

The first metric of interest was the use of automation scripts. A typical automation script ("pre-pass[2]") will issue tens to hundreds of directives needed to configure a tracking pass. Automation scripts are not

[2] The DSN uses three key automation scripts, of TDNs (Temporal Dependency Network). The pre-pass TDN is used to configure all the equipment and start the pass, including all signal acquisition. It is the most

always an option – if there is an equipment failure, or there is a last-minute change in the plan, the operator has to resort to manual steps. The metric we selected was the percent of tracking passes that use automation scripts, we set a target goal of 100% for the three common automation scripts – pre-pass, in-pass, and post-pass. Figure 1 shows the values for this metric at the beginning of this exercise. What we found that::

1. The use of the automation scripts is not consistent across the antenna complexes

2. The use of automation scripts is not consistent among the three scripts

3. The percent-used was at times significantly below 100%

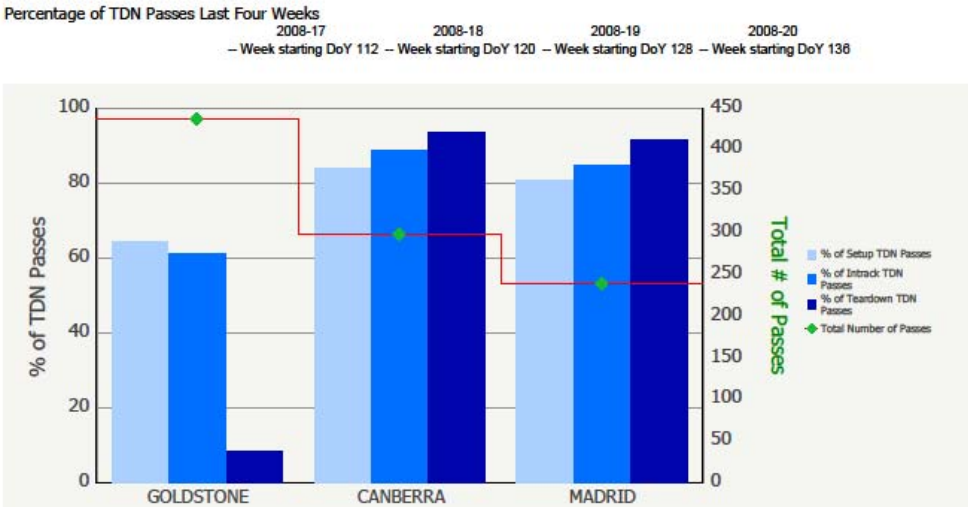


Figure 1 – Use-percentage of Automation Scripts (pre-BI)

Figure 2 – Use-percentage of Automation Scripts (With BI feedback)

Once the metrics were distributed and discussed, it took a modest effort (programmatically and technically) to create a dramatic improvement as shown in Figure 2 – better consistency across the complexes as well as across the three scripts and higher percentage-used. Much of the change was self-directed – BI metrics have that effect.

---

complex TDN. The in-track TDN is used to monitor the tracking pass and correct minor anomalies. The post-pass script TDN closes the pass and restores the equipment to nominal position.

And only a modest effort is required to keep the results going – the BI metrics are issued weekly to management and interested staff, a powerful reminder.

IV. Delivery 2 – Adding Data Types to the M&C Database

At times, the existing data stores are not sufficient to provide the desired metrics. This can be solved by one of two methods:

1. If the desired data is electronically created, but just not stored, the solution would be to modify the underlying software, with no change to the enterprise operations processes, to store the desired data.

2. If the desired data is not created electronically, the solution will impact the enterprise operation and should be carefully designed to maximize cost effectiveness.

In the DSN example, when we started probing the reasons for the less-than-100% use of automation scripts, we found a mix of factors that could be measured, such as equipment failures, and factors that could not be measured. The prime example of the latter was non-electronic communications between the DSN and the missions. There clearly were incidents where last-minute phone conversations with the missions caused the operator to abort an automation script and resort to manual control. But there was no electronic record of these events, let alone how often they occurred.

Deciding how to proceed is a delicate balance between the cost and the benefit of collecting this data. We have settled on a collection mechanism with modest impact: an electronic survey that is required to be completed by the operator at the completion of each pass requiring only yes/no check marks. We needed to balance the extra effort on the users (primarily training and the time to fill out the surveys) and the benefit.

With immense data stores, BI carries the risk of producing metrics, and lots of them, with minimal effort. For an effective use of BI, it is imperative to ask early on "What will I do with this data and what is the target audience?" In this case we have estimated early on (but had no apriory proof) that a significant portion of the non-electronic communications between the DSN and the missions were not of tangible added value and could be reduced with a resulting cost savings to the DSN and to the missions.

The survey measures five types of non-electronic communications: two types of change in mission parameters and three types of phone communications with no change to mission parameters. Figures 3 shows the results from this survey, aggregated across all the missions – similar charts show the results for a specific mission. With these reports we have started the process and agreeing with each mission on what non-electronic communications is, or is not needed, given the cost impact of such communications.

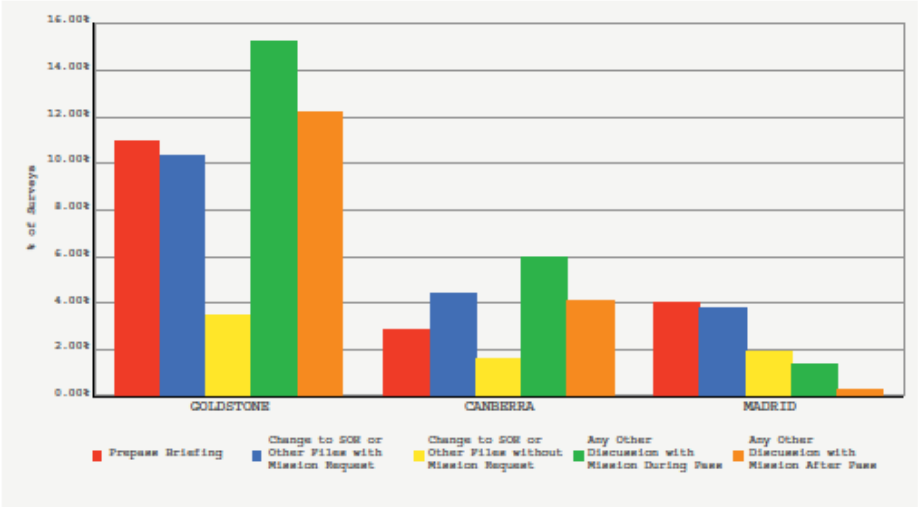
V. Delivery 3 – Affecting Areas beyond M&C

In Sections III and IV, BI extracted from the M&C database was used to improve the cost efficiency of M&C operations. But the data in the M&C database can be mined to improve the efficiency of the enterprise in other areas as well. A smart deployment of BI tools will require experts from different areas of the enterprise to maximize efficient use.

In the DSN example, we observed that the M&C database contains a large number of "Event messages"[3] send to the operator, as well as those sent directly to the database without being displayed to the operator (log only). Thus, mining that data could provide leads to the maintenance teams, indicating what hardware and software have failed, or require maintenance attention.

[3] The primary "Event Message" is a warning. In this paper we will use the term "warning" as a catch-all to all event messages

2010-01 -- Week starting DoY 004 to 2010-04 -- Week starting DoY 025



DSCC Location	Prepass Briefing	Change to SOE or Other Files with Mission Request	Change to SOE or Other Files without Mission Request	Any Other Discussion with Mission During Pass	Any Other Discussion with Mission After Pass
GOLDSTONE	10.90%	10.26%	3.42%	15.17%	12.18%
CANBERRA	2.80%	4.36%	1.56%	5.92%	4.05%
MADRID	3.99%	3.72%	1.86%	1.33%	0.27%

Figure 3 – Non-electronic Interfaces

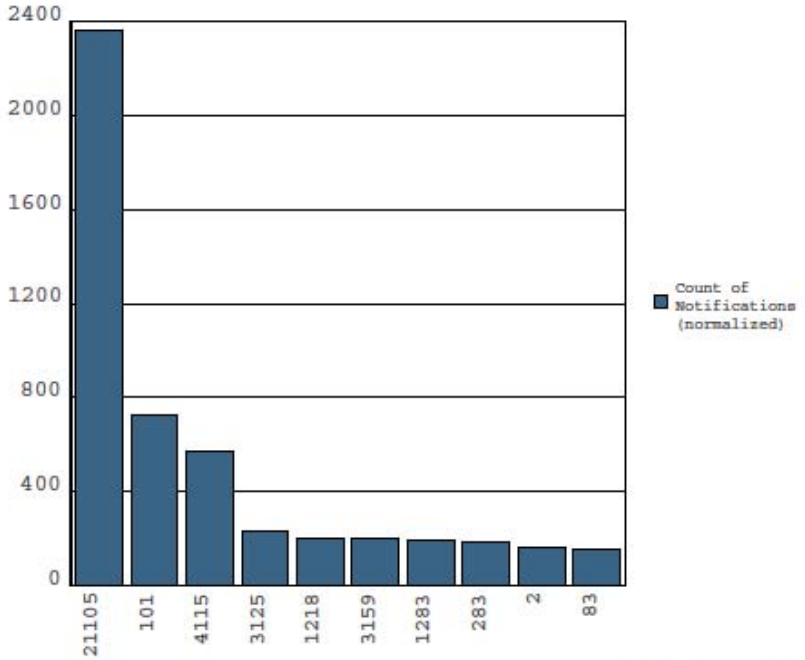
So the issue became the design of metric(s) that will be useful to the maintenance teams, without overwhelming them with less-useful metrics. The problem was compounded by the fact that the event messages were developed over multiple generations of equipment and was of varying fidelity.

The report we selected, shown in Figure 4, is to have charts of “top-10”, listing in histogram and table formats the “top-10” alarms at each complex. Human analysis would then determine whether the solution is a maintenance activity (e.g. calibrating a noise diode or adjusting an operational table) or a software solution (to improve the fidelity of the warning).

Top 10 Event Notifications for Last 4 Weeks

2010-01 -- Week starting DoY 004 01 2010	2010-02 -- Week starting DoY 011 02 2010	2010-03 -- Week starting DoY 018 03 2010	2010-04 -- Week starting DoY 025 04 2010

GOLDSTONE



CIMS Assembly		Event Type	Event ID	Metrics	Count of Notifications (normalized)	Total Number of Notifications (normalized)
UCS	Uplink Consolidated Software	C!	21105		2,359	35,392
WVSR	Wideband VLBI Science Receiver Software	CA	101		726	2,903
DCC	Downlink Channel Control Software	W!	4115		574	1,148
DCC	Downlink Channel Control Software	W!	3125		232	3,487
CE	Connection Engines	PA	1218		204	407
DCC	Downlink Channel Control Software	W!	3159		201	2,816
CE	Connection Engines	W!	1283		196	20,428
CE	Connection Engines	W!	283		188	19,559
CE	Connection Engines	PA	2		161	74,602
FSPA	Full Spectrum Array Software	PA	83		156	5,301

Figure 4 – Top-10 Event Messages

VI. Conclusions

Our experience in using BI tools was very positive. With these tools we were able to reduce the cost of preparing standard reports, create new insights to the operations of the systems and enabling improvement in its cost efficiency by , allowing a smaller team to create equal or better performance. A key to that efficiency is that the BI tools are created, and amortized over a large set of enterprises, reducing the development cost.